Title: CBP-IDRs regulate acetylation and gene expression

Authors

Katie L. Gelder[1,4,5], Nicola A. Carruthers[1,4], Grace Gilbert[1,4,6*], Laura Harrison[1,4*], Thomas Evans[1], Sophie S. Ball[1,8], Mark Dunning[7], Timothy D. Craggs[3,4,5], Alison E. Twelvetrees[2,4,5,6], Daniel A. Bose[1,4,5,6].

Affiliations

1. Molecular and Cellular Biology, School of Biosciences, Faculty of Science, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK.
2. Division of Neuroscience, The School of Medicine and Population Health, Faculty of Health, The University of Sheffield, Glossop Road, Sheffield, S10 2HQ, UK.
3. Department of Chemistry, School of Mathematical and Physical Sciences, Faculty of Science, The University of Sheffield, Dainton Building, Brook Hill, S3 7HF, UK.
4. Single Molecules at Sheffield (SM@SH), The University of Sheffield.
5. Institute for Nucleic Acids, The University of Sheffield.
6. Neuroscience Institute, The University of Sheffield.
7. Sheffield Bioinformatics Core, Faculty of Health, The University of Sheffield.
8. Department of Immunology and Inflammation, Imperial College London.

*Authors contributed equally to this work.

Corresponding author: Daniel Bose, e-mail: d.bose@sheffield.ac.uk

Keywords

CBP/p300, gene regulation, biomolecular condensates, transcription, IDRs, Intrinsically disordered proteins, enhancers, acetylation.

Abstract

Intrinsically disordered regions (IDRs) have emerged as crucial regulators of protein function, allowing proteins to sense and respond to their environment. *Creb binding protein* (CBP) and *EP300* (p300) are transcription coactivators that regulate gene expression in multicellular organisms, following their recruitment to *cis*-regulatory elements. CBP and p300 contain large IDRs, however little is known about how these different IDRs work together to regulate CBP function. Here, we show that CBP-IDRs cooperate to control different aspects of CBP behaviour in the nucleus, by regulating the properties of fluid-like condensates formed by endogenous CBP. We show how IDRs with different sequence properties make unique contributions to CBP behaviour by establishing a balance between positive and negative regulation of CBP condensates. These conflicting interactions are functionally important, shaping CBPs response to factors such as lysine acetylation, that influence condensate formation. When disrupted, regulatory CBP-IDRs change how CBP interacts with chromatin, alter patterns of CBP-dependent histone acetylation and change gene expression. Together, our work highlights how IDRs with different sequences, spatially segregated in the same protein, can cooperate to shape protein function.

Introduction

Intrinsically Disordered Regions (IDRs) are amino acid sequences that regulate many biological functions despite lacking stable secondary and tertiary structures(van der Lee et al. 2014; Holehouse and Kragelund 2023). IDRs mediate specific and non-specific interactions between biomolecules and control the formation and behaviour of biomolecular condensates. Functionally important regions in IDRs can be defined by patterns of amino acid properties or compositional biases[1,2]. Their ability to adopt multiple conformations means that IDRs are responsive to their interactions and environmental context[3,4]. Despite an increasing understanding of the rules which govern these processes, how the different properties of IDRs are combined in the same protein to generate specific functions is still poorly understood.

Creb binding protein (CBP) and its paralog *EP300* (p300) are histone acetyltransferases (HATs) essential for shaping gene expression from *cis*-regulatory elements (CREs). By combining a structured catalytic HAT domain with large IDRs, CBP and p300 integrate HAT activity with the conformational plasticity needed to interact with diverse transcription factors (TFs) at CREs. CBP and p300 control the strength of gene transcription by finely balancing lysine acetylation on histones (e.g H3K27ac, H3K18ac and H3K122ac)[5–8] and transcription factors (TFs)[9], contribute to RNA polymerase II (PolII) recruitment to maintain promoter-proximal pausing[10–12], and form condensates to regulate transcriptional bursting[11,13,14]. IDRs and condensates are central to many of these processes. Firstly, binding to TFs is mediated by a combination of structured binding domains and IDRs, many of which undergo 'coupled folding and binding' with TFs [15–20]. Secondly, HAT activity is regulated by the autoregulatory loop (AL), an IDR insertion in the HAT domain that regulates enzymatic activity by blocking substrate binding to the active site[21]. The AL undergoes autoacetylation and binds to non-coding enhancer RNAs (eRNAs), both of which displace it from the active site to stimulate activity [21–25] [21,22,26]. p300 HAT activity is decreased by crowding reagents and when condensate formation is triggered *in vitro* [27,28]. Conversely, formation of CBP and p300 condensates driven by the disordered transactivation domains of TFs in cells increased acetylation across the genome[11], while TF binding to regions outside the HAT domain in CBP have been shown to stimulate or inhibit its HAT activity[29]. In combination, this suggests that CBP functions are highly dependent on interactions with its environment, and underlines the importance of IDRs for this process.

Understanding how multiple IDRs with different sequence properties regulate functions when combined in the same protein remains a significant challenge. This is particularly relevant for CBP and p300, where multiple IDRs mediate diverse biological processes [11,19,30]. To address this question, we examined how distinct IDRs within CBP (CBP-IDRs) work together to regulate CBP behaviour, by systematically characterising the impact of these IDRs on CBP's function within cells. We found that different IDRs play very different roles. Two adjacent C-terminal IDRs with different amino acid compositions together allow CBP to respond to other factors that regulate condensates, such as acetylation. When disrupted, these IDRs caused changes in the behaviour of CBP condensates, and altered CBP's chromatin localisation, histone acetylation and gene

expression. Our results uncover new and unexpected functional roles for CBP-IDRs with different sequence properties, and highlight how intramolecular cooperation between IDRs can shape protein behaviour.

Results

CBP forms fluid-like condensates in the nucleus.

To validate the formation of CBP condensates, we imaged endogenous CBP by immunofluorescence (IF) in HEK293T cells, revealing distinct nuclear puncta consistent with previous observations[31–34] (Figure 1A). To uncover the behaviour of these puncta, we introduced an in-frame HaloTag to the C-terminus of endogenous CBP (CBP-Halo). PCR from genomic DNA confirmed heterozygous insertion (Figure S1A-B) and we confirmed near-endogenous expression levels by RT-qPCR (Figure S1C) and western blotting (Figure 1B). To confirm that CBP-Halo can also form endogenous puncta, we carried out live-cell imaging using the fluorescent ligand TMR to label the HaloTag (CBP-Halo$_{TMR}$). CBP-Halo$_{TMR}$ puncta were only observable in CBP-Halo HEK293T cells upon addition of TMR (Figure 1C).

Next, we used lattice light sheet microscopy to visualise the spatial organisation of CBP puncta (Figure 1D). CBP-Halo formed an average of 27.9 condensates/nucleus and we observed a variation in the size and integrated intensity density (IID) of observable puncta, with a mean volume of 33 ± 14.6 and an IID of 835.6 ± 132.8. Importantly, CBP-Halo$_{TMR}$ condensates demonstrated hallmarks of fluid-like behaviour, as they were able to fuse and disperse over the 2s imaging period in live-cell confocal microscopy (Figure 1E, S1D & Supplementary Movie 1). We next carried out Fluorescence Recovery After Photobleaching (FRAP) on endogenous CBP using Janelia Fluor® 549 Halo ligand (CBP-Halo$_{JF549}$), which showed that fluorescent signal recovered to ~50% its pre-bleach levels after 30s, and to ~90% over a 60s timeframe following photobleaching (Figure 1F-G & Supplementary Movie 2). Together, these data clearly establish that endogenous CBP condensates have fluid-like behaviour in their native nuclear environment.

IDRs in CBP have distinct abilities to form condensates.

Full length CBP$_{wt}$ is predicted to be ~65% disordered (Figure 1H& S1E)[19]. AlphaFold2 predictions of CBP$_{wt}$ show structured domains interspersed with large disordered regions (Figure S1F)[35]. To test how these regions regulate CBP's function, we identified seven IDRs for further analysis, using a threshold of 50 consecutive amino acids with a predicted disorder >50% (CBP$_{IDR1-IDR7}$, Figure 1H & S1E)[36] in PONDR[37] to define IDR boundaries. We also tested the CBP:FUS interaction domain (residues 1892-2441, CBP$_{CFID}$), a region that bound to Fused-in-sarcoma (FUS) [11,38–41] and which comprises two identified IDRs (CBP$_{IDR6}$ and CBP$_{IDR7}$) and the Nuclear coactivator binding domain (NCBD), a TF binding region[15,17]. Finally, we included the Autoregulatory Loop (AL, residues 1558-1607), an IDR important for regulating CBP activity via acetylation

and RNA binding (Figure 1H)[21,22,28,42]. The selected IDRs are characterised by distinct amino acid compositions (Figure S1G) and sequence properties, including different values for their fraction of charged residues (FCR) and net charge per residue (NCPR) [43] (Figure S1I). As these factors affect biomolecular condensates [44], we tested the ability of each CBP-IDR to form condensates in HEK293T cells using the OptoDroplet system[36,39]. Candidate IDRs and controls were fused to the photoactivatable Photolyase homology region of Cry2 (Cry2PHR) and an SV40 nuclear localization sequence (NLS) to control phase transitions in the nucleus (Figure S1I). Following exposure to blue light, CBP-IDRs with strong condensate forming properties (e.g CBP_{IDR7}) formed defined puncta, while IDRs without defined condensate forming properties (e.g CBP_{IDR6}) failed to form clear puncta (Figure 1I). To quantify differences between CBP-IDRs, we measured the standard deviation of treated nuclei: as condensate formation causes strong intensity changes, we expected the variance in grey values for a defined nucleus to increase, resulting in a high standard deviation. Notably, most tested CBP-IDRs behaved similarly to the negative control, with the exception of CBP_{IDR7} and CBP_{CFID} which both demonstrated a rapid increase in standard deviation following photoactivation (Figure 1J). We then calculated the initial rate of droplet formation over the first 30 seconds following photoactivation; again, only CBP_{IDR7} and CBP_{CFID} showed a greater ability to form condensates than the negative control Opto-NLS (Figure 1K). The results suggest that different CBP-IDRs undergo distinct phase transitions in the nucleus, and highlight strong condensate forming properties in the disordered C-terminal domain of CBP (CBP_{IDR7} and CBP_{CFID}).

CBP-IDRs regulate the number of CBP condensates

We next tested how CBP-IDRs contribute to formation of full-length CBP (CBP_{wt}) condensates. We deleted identified CBP-IDRs in CBP_{wt} fused to an in-frame C-terminal GFP-tag (CBP_{wt}-GFP, Figure 2A). To generate $CBP_{\Delta AL}$-GFP, we replaced the IDR with a flexible 12 residue linker to maintain correct folding of the HAT domain (Figure 2A) [23]. HEK293T cells expressing the CBP-GFP constructs showed similar levels of CBP-GFP (Figure S2A). Cells were then fixed and imaged by confocal microscopy for quantification of nuclear puncta (Figure 2B).

Data were initially selected for successful transfection based on the presence or absence of GFP signal, then the data was blinded and scored for punctate or diffuse nuclear signal (Figure 2C). Following expression of CBP_{wt}-GFP, $CBP_{\Delta IDR1}$-GFP or $CBP_{\Delta IDR6}$-GFP, 100% of scored nuclei demonstrated a punctate GFP signal. Indeed, expression of the majority of CBP-GFP constructs resulted in punctate GFP signal, as $CBP_{\Delta IDR3}$-GFP, $CBP_{\Delta IDR4}$-GFP and $CBP_{\Delta AL}$-GFP were all >91% punctate. $CBP_{\Delta IDR7}$-GFP and CBP_{CFID}-GFP were exceptions, with only 52.9% and 40.0% of nuclei showing a punctate signal. However, this coarse-grained approach was insufficient to describe the full spectrum of condensate-forming $CBP_{\Delta IDR}$-GFP behaviour.

To examine these behavioural differences in more detail, we applied an unbiased filter to remove oversaturated nuclei, noise and overlapping puncta (Figure S2B,C&D) and calculated the mean number of condensates per

nucleus for each CBP-GFP construct (Figure 2D). CBP_{wt}-GFP formed well defined condensates (average 21.1 ± 2.8 s.e.m condensates/nucleus, Figure 3B&D). The majority of CBP-IDR deletions (except $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP) also formed condensates, as >80% of nuclei displayed a punctate signal and nuclei that passed the filter formed >6 condensates per nucleus (Figure 2D). $CBP_{\Delta IDR1}$-GFP (22.8 ± 6.8 s.e.m condensates/nucleus) and $CBP_{\Delta IDR3}$-GFP (19.9 ± 3.7 s.e.m condensates/nucleus) behaved similarly to CBP_{wt}-GFP, while $CBP_{\Delta IDR4}$-GFP (9.8 ± 2.2 s.e.m condensates/nucleus) and $CBP_{\Delta AL}$-GFP (10.9 ± 1.2 s.e.m condensates/nucleus) showed a ~50% decrease in the average number of condensates compared to CBP_{wt}-GFP (Figure 2D). 60% of $CBP_{\Delta CFID}$-GFP and 47.1% of $CBP_{\Delta IDR7}$-GFP nuclei displayed a diffuse GFP signal (Figure 2C) and these cells also showed a ~75% decrease in quantifiable condensates per nucleus compared to CBP_{wt}-GFP ($CBP_{\Delta CFID}$-GFP, 5.6 ± 1.4 s.e.m condensates/nucleus; $CBP_{\Delta IDR7}$-GFP, 5.3 ± 1.1 s.e.m condensates/nucleus (Figure 2D). Visual inspection of $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP nuclei showed that unlike other CBP-GFP constructs, any residual condensates which passed filtering were poorly defined puncta (Figure 2B). We were therefore confident that both $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP failed to form visible condensates. Together with the OptoDroplet results (Figure 1 J,K), this emphasised the importance of C-terminal IDRs (CBP_{CFID} and CBP_{IDR7}) for promoting CBP_{wt} condensates. However, in contrast to other CBP-IDRs, $CBP_{\Delta IDR6}$-GFP formed ~30% more condensates (27.6 ± 1.9 s.e.m condensates/nucleus) than CBP_{wt}-GFP (Figure 2D), a surprising result as CBP_{IDR6} constitutes a large part of $CBP_{CFID,}$ which blocked condensates when deleted.

The CBP-IDR deletions also affected the variation in the numbers of condensates. For CBP_{wt}-GFP, the standard deviation (SD) in the number of condensates was 0.86 of the mean (SD = 18.3, mean = 21.1). The majority of CBP-IDR deletions displayed a similarly high ratio of SD to mean, ranging from 1.37 for $CBP_{\Delta IDR1}$-GFP to 0.75 for $CBP_{\Delta IDR7}$-GFP. The exception was again $CBP_{\Delta IDR6}$-GFP, where the SD was only 0.49 of the mean (SD = 13.6, mean = 27.6; Figure 3D). The result suggests that different IDRs can affect the variability in the number of condensates as well as the overall number of condensates.

CBP-IDRs regulate the size and intensity of CBP condensates

Deletion of CBP-IDRs altered the size and intensity of CBP condensates (Figure 2B). To quantify this behaviour, we calculated the IID of puncta for condensate-forming $CBP_{\Delta IDR}$-GFP constructs (Figure 2E). CBP_{wt}-GFP condensates had relatively uniform size and intensity (mean IID = 870 ± 72 s.e.m.), and most $CBP_{\Delta IDR}$-GFP constructs again behaved similarly to CBP_{wt}-GFP. The two exceptions were $CBP_{\Delta AL}$-GFP which demonstrated a 1.5-fold increase and $CBP_{\Delta IDR6}$-GFP which demonstrated a 3-fold increase in integrated density compared to CBP_{wt}-GFP (Figure 2E). The results suggest that CBP_{IDR6} and CBP_{AL} may restrict the size and intensity of full length CBP condensates.

To characterise the properties of the larger CBP-GFP condensates, we carried out FRAP on individual CBP_{wt}-GFP, $CBP_{\Delta AL}$-GFP and $CBP_{\Delta IDR6}$-GFP puncta (Figure 2F-G). Following bleaching, all condensates

recovered rapidly (Recovery rate: CBP_{wt}-GFP = 0.11 s^{-1}, $CBP_{\Delta AL}$-GFP = 0.12 s^{-1}, $CBP_{\Delta IDR6}$-GFP = 0.13 s^{-1}, Figure 2G, S2E), thus demonstrating fluid-like behaviour rather than solid or gel-like properties, and only subtle differences in nuclear mobility.

We also treated cells transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP with 5% 1,6-hexanediol (1,6-HD), an aliphatic alcohol that disrupts weak hydrophobic interactions[45]. Condensates formed by all three constructs dissolved in <20s following application of 5% 1,6-HD (Figure 2H), suggesting that hydrophobic interactions are important for maintaining CBP condensates.

Diverse RNAs are important for shaping condensate properties. As CBP binds to RNA *in vitro* and in cells, we therefore tested how it affects CBP condensates. Treating cells with 0.5uM RNAse A decreased the proportion of punctate nuclei for all CBP-GFP constructs (Figure S2F). Notably, the size and intensity of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP condensates was decreased following addition of RNase A, but residual $CBP_{\Delta IDR7}$-GFP condensates were unchanged (Figure 2I-J). RNase A also caused an increase in the number of condensates formed for all tested constructs (Figure S2G), potentially reflecting a dispersal of condensates following RNA degradation. The results highlight the ability of undefined RNAs to maintain properties of CBP condensates.

Taken together, these results highlight clear differences between condensate behaviour regulated by CBP-IDRs. Notably, CBP_{IDR7} strongly promotes, while CBP_{IDR6} restricts CBP condensates. CBP_{AL} and CBP_{IDR6} play a more complex regulatory role, restricting the size and intensity of condensates, with CBP_{IDR6} in particular affecting variability in the number of condensates formed. Large CBP condensates have fluid-like properties and are sensitive to disruption of hydrophobic contacts, and RNAs can help to promote CBP condensate formation.

CBP IDRs regulate endogenous CBP condensates

Endogenous CBP-Halo (Figure 1E-G) and CBP-GFP (Figure 2F-G) both form fluid-like condensates. We therefore asked how CBP-IDRs affect behaviour of endogenous CBP. We used lattice light sheet microscopy to image endogenous CBP-Halo labelled with fluorescent JF_{646} ligand (CBP-$Halo_{JF646}$) following expression of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP (Figure 2K). Compared to control cells CBP_{wt}-GFP caused a slight decrease in the number of endogenous CBP-Halo condensates (27.9 ± 79.2 > 14.9 ± 28.9 s.e.m condensates/nucleus; Figure 2L), coupled to an increase in the IID (835 ± 132 > 1759 ± 803 s.e.m; Figure 2M) and average volume (32.9 ± 14.6 > 58.9 ± 40.0 s.e.m; Figure 2N), probably reflecting the presence of extra CBP-GFP molecules due to transient expression. However, compared to cells expressing CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP increased the median IID (2223 ± 832 s.e.m; Figure 2M) and the volume (101.7 ± 55.4 s.e.m; Figure 2N) of endogenous CBP-Halo condensates. The opposite was seen for $CBP_{\Delta IDR7}$-GFP, which decreased both the IID (655 ± 411 s.e.m; Figure 2M) and volume (45.2 ± 33.4 s.e.m; Figure 2N) of endogenous CBP-Halo condensates. Similar results were observed by live cell confocal microscopy (Figure S2H) with no bleed through into other imaging channels (Figure S2I). Although we cannot rule out complexities arising

from the mixed population, these results show that CBP-IDRs can shape the behaviour of endogenous CBP condensates, exerting a dominant negative/positive influence on endogenous CBP behaviour.

Opposing properties of CBP-IDRs reflect their sequence characteristics

To investigate whether the opposing behaviours of the C-terminal IDRs CBP_{IDR6} and CBP_{IDR7} could be derived from their amino acid sequence, we used localCIDER[43] to examine their sequence properties of alongside CBP_{CFID} which contains both CBP_{IDR6} and CBP_{IDR7} (Figure 3A). Compared to CBP_{wt}, CBP_{CFID} is enriched for polar amino acids and proline, which promote expanded chain conformations[46], but depleted for charged and aromatic residues. There is more variation between CBP_{IDR6} and CBP_{IDR7} individually: Polar amino acids are more enriched in CBP_{IDR7} (54%) than CBP_{IDR6} (41%), while CBP_{IDR6} is enriched for proline (21%) compared to CBP_{IDR7} (14%). This distribution also causes CBP_{IDR6} to have a higher FCR and NCPR than CBP_{IDR7} (Figure S3A). The linear distribution of amino acids (Figure 3B) shows that amino acid properties in CBP_{IDR6} are evenly distributed, but less so in CBP_{IDR7}. This is especially true for polar amino acids, which in CBP_{IDR7} are concentrated around an 18-residue polyQ region (residues 2199-2216), and proline which is concentrated towards the C-terminus (Figure 3B).

Because of this uneven distribution, it's possible that CBP_{IDR6} and CBP_{IDR7} possess distinct clustering of amino acid properties, a sequence feature that can modulate condensate behaviour. The parameter κ quantifies charged residue clustering in IDRs (0=mixed, 1=clustered)[43,47], and the parameter Ω measures proline/charged residue mixing (0=mixed, 1=clustered)[46]. In both cases, low values favour extended conformations. C-terminal CBP-IDRs have low values of κ (CBP_{IDR6}=0.187, CBP_{IDR7}=0.206; Figure S3A), indicating well-mixed charges and extended chains. Similarly, CBP_{CFID} has relatively low Ω (Ω=0.278) indicative of relatively little clustering of proline/charged residues across the whole C-terminus. However, this masks differences between CBP_{IDR6} and CBP_{IDR7}; Ω is low for CBP_{IDR6} (Ω=0.161), while in CBP_{IDR7} Ω is much higher (Ω=0.385) (Figure 4L), indicating a difference in their proline/charged residue distribution and suggesting that CBP_{IDR6} and CBP_{IDR7} could adopt different conformations.

To test the role of sequence patterning in CBP_{IDR6} and CBP_{IDR7}, we used GOOSE[48] to design CBP-IDR mutants where the overall composition of CBP_{IDR6} and CBP_{IDR7} was retained, but their sequence was randomly shuffled to disrupt their sequence patterning ($CBP_{IDR6\text{-}Shuffle}$ and $CBP_{IDR7\text{-}Shuffle}$ Figure S3B,C). When imaged using confocal microscopy (Figure 3C) and following filtering (Figure S3D), $CBP_{IDR6\text{-}Shuffle}$-GFP had a large effect similar to $CBP_{\Delta IDR6}$-GFP on CBP condensates, increasing the proportion of punctate nuclei (Figure S3E), the number of condensates per nuclei (Figure 3D) and the size and intensity of condensates (Figure 3E), while the variability in the number of condensates was decreased. The result suggests that the sequence of amino acids in CBP_{IDR6} is important for its behaviour. However, $CBP_{IDR7\text{-}Shuffle}$-GFP behaved comparably to CBP_{wt}-GFP, with the proportion of punctate nuclei (Figure S3E), number of condensates per nuclei (Figure 3D) and the

size and intensity of condensates (Figure 3E) all alike, indicating that the behaviour of CBP_{IDR7} is more a function of its overall composition than of a defined amino acid sequence.

Curiously, despite disrupting the sequence of amino acids, $CBP_{IDR6\text{-}Shuffle}$ and $CBP_{IDR7\text{-}Shuffle}$ actually displayed very little difference in the calculated values for κ and Ω (Figure S3A, CBP_{IDR6}: κ=0.187, Ω =0.161; CBP_{IDR7} κ=0.214, Ω=0.393) compared to CBP_{wt}. This similarity in patterning is consistent with the equivalent behaviour of $CBP_{IDR7\text{-}Shuffle}$-GFP and CBP_{wt}-GFP. However, the similarity in patterning parameters but difference in behaviour between $CBP_{IDR6\text{-}Shuffle}$-GFP and CBP_{wt}-GFP, suggests that the contrasting behaviour of $CBP_{IDR6\text{-}Shuffle}$-GFP is not a direct result of charged residue or proline/charged residue clustering.

Next, as enrichment for polar residues is associated with prion-forming domains[49], we examined the sequence of the C-terminal CBP-IDRs using PLAAC analysis[49,50] which predicts prion-like regions from amino acid sequences. PLAAC analysis predicted that in CBP_{wt}, CBP_{IDR6} forms two discontinuous prion-like regions, and that CBP_{IDR7} forms a large, contiguous prion-like domain (Figure 3F). These predicted properties were disrupted in both $CBP_{IDR6\text{-}Shuffle}$, whose predicted prion-like behaviour diverged from CBP_{wt}, and in $CBP_{IDR7\text{-}Shuffle}$ where prion-like properties were predicted to extend further towards the C-terminus of CBP, potentially caused by the shuffling of prolines away from this region (Figure 3F).

Finally, we noted that residues 1951-1978 in CBP_{IDR6} ($CBP_{IDR6\text{-}H2}$), which were shown to form transient α-helices by Nuclear Magnetic Resonance (NMR) [51], and were predicted by Alphafold to interact with an α-helix formed by residues 2187-2216 in CBP_{IDR7} (Figure S3F,G,$CBP_{IDR7\text{-}H1}$) did not affect CBP condensates when mutated (Figure S3H-J).

Taken together, the data highlight differences in the sequence properties of CBP_{IDR6} and CBP_{IDR7}, particularly the relative abundance of polar residues and proline, clustering of proline/charged residues and predicted prion-like properties. These could account for their divergent behaviour in OptoDroplet experiments (Figure 1) and when deleted in CBP-GFP (Figure 2). Although we could not confidently link changes in behaviour to changes in the clustering of amino acid properties (κ and Ω) in CBP-IDR shuffle mutants, $CBP_{IDR6\text{-}Shuffle}$ caused large changes in CBP condensates, similar to $CBP_{\Delta IDR6}$-GFP, suggesting that its amino acid sequence is important for regulating CBP-condensates.

The different behaviours caused by CBP_{IDR6} and CBP_{IDR7} led us to ask whether their sequence properties were conserved in CBP's paralog p300. Over their entire sequence, CBP and p300 are 60.8% identical (70.7% similarity). While this identity is as high as 87.3% (94.4% similarity) in the HAT domain, CBP_{IDR6} only has 48.1% identity (58.0% similarity) and CBP_{IDR7} has 56.1% identity (66.2% similarity) (Figure S3K). Interestingly, this difference is reflected in PLAAC analysis[49,50] where $p300_{IDR6}$ forms a single, contiguous prion-like domain in contrast to the disrupted prion-like properties of CBP_{IDR6} (Figure S3L).

CBP-IDR$_6$ allows acetylation to shape CBP condensates.

Our results show that CBP-IDRs cause condensates to exhibit very different properties. We reasoned that this variability could help other environmental inputs, such as PTMs, to shape condensate behaviour and affect CBP function. To test this, we focused on acetylation, which can regulate p300 condensates in cells by providing nucleation sites for bromodomain proteins[11,28,52] and *in vitro*, where p300 HAT domain condensates are promoted by the AL's positive charge and negatively regulated by lysine acetylation [28].

Firstly, we specifically depleted CBP and p300-dependent acetylation using A-485, a potent inhibitor of CBP and p300 HAT activity[53] (Figure 4A). Treatment with 500nM A-485 for 36 hours[53] increased the proportion of punctate nuclei in CBP$_{wt}$-GFP (Figure 4B), but had no effect on the number (Figure 4C) or IID of condensates (Figure 4D). In contrast, there was no change in the proportion of CBP$_{\Delta IDR6}$-GFP cells forming condensates (Figure 4B), a modest decrease in the number of puncta formed per nuclei (Figure 4C), but a large decrease in the size and intensity of condensates (Figure 4D). For CBP$_{\Delta IDR7}$-GFP, there was a large increase in the proportion of punctate nuclei (Figure 4B), but when quantified, the size and intensity of condensates was lower than control cells (Figure 4D).

We next induced hyperacetylation and chromatin de-compaction by treating cells transfected with CBP-IDR mutants with Trichostatin A (TSA), a general inhibitor of histone deacetylases (HDACs). Again, treatment with 500nM Trichostatin A (TSA)[54] for 90 minutes had a negligible effect on CBP$_{wt}$-GFP and CBP$_{\Delta IDR7}$-GFP cells. However, CBP$_{\Delta IDR6}$-GFP showed an increase in the number of condensates per nuclei and the size and intensity of condensates formed (Figure 4B-D), potentially indicating an increased dignity to acetylation following removal of CBP$_{IDR6}$.

We were surprised by the A-485 data, as A-485 had previously been used to cause a large increase in p300 puncta [11]. We therefore tested how CBP condensates responded to different concentrations (including 10uM as used previously)[11] of A-485 (Figure 4E). CBP$_{wt}$-GFP and CBP$_{\Delta IDR6}$-GFP showed a slight increase in the median number of condensates formed per nucleus, which was more pronounced after 2 hours treatment and with 10um A-485 (Figure 4F & S4D). The effect of A-485 on the size and intensity of condensates was more evident; both CBP$_{wt}$-GFP and CBP$_{\Delta IDR6}$-GFP showed a decrease in size and intensity of condensates across all tested A-485 concentrations (Figure 4G). However, after 2 hours treatment, CBP$_{wt}$-GFP showed no significant difference between control cells and those treated with A-485, while the size and intensity of condensates was still significantly decreased for CBP$_{\Delta IDR6}$-GFP (Figure S4E). Given the limited changes in size and intensity of CBP$_{wt}$-GFP condensates after 48 hours of A-485 treatment, compared to the significant decrease in intensity observed in CBP$_{\Delta IDR6}$-GFP condensates, these findings suggest that the absence of CBP$_{IDR6}$ might render condensates more sensitive to acetylation regulated by CBP and p300. We also note the slight recovery of CBP$_{wt}$-GFP condensate size and intensity at 10uM A-485, compared to 1uM (Figure 4G) This potentially reflects the different roles played by direct acetylation of CBP (downregulates condensates) and indirect acetylation of factors that nucleate CBP binding (upregulates condensates). This effect was also not evident for CBP$_{\Delta IDR6}$-GFP, further supporting a role for CBP$_{IDR6}$ in controlling sensitivity to acetylation.

Next, to assess the role of direct acetylation on CBP condensates, we examined in more detail how the AL regulates CBP condensates. In $CBP_{\Delta AL}$-GFP data, we observed a ~50% reduction in the number of condensates (Figure 2D) but an increase in their size and intensity (Figure 2E) compared to CBP_{wt}-GFP. This suggested that CBP condensates could be balanced by positive contributions due to exposed lysines in the AL and negative regulation due to lysine acetylation. To test this hypothesis, we generated CBP-AL mutations, $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP, where all 13 lysine residues in the AL were mutated to glycine to prevent acetylation (Figure S4F), on a background of CBP-IDR deletion mutants, and expressed them in HEK293T cells (Figure 4H & S4G). On a CBP_{wt}-GFP background, $CBP_{wt/\Delta AL}$-GFP and $CBP_{wt/KTG}$-GFP had similar effects, decreasing the overall proportion of punctate nuclei compared to CBP_{wt}-GFP (Figure 4I). Both CBP-AL mutations also reduced the number of condensates per nucleus by 40-60% compared to CBP_{wt}-GFP, but retained a high variability in the number of condensates (Figure 4J), and caused a 40-50% increase in the size and intensity (Figure 4K). As the AL in the $CBP_{wt/KTG}$-GFP construct was intact, but incapable of being acetylated, this suggests that acetylation of lysine residues in the AL can negatively regulate the size and intensity of full length CBP condensates on a CBP_{wt}-GFP background in cells, similar to the behaviour of p300-HAT *in vitro*[28].

Next, we tested how C-terminal IDRs (CBP_{IDR6} and CBP_{IDR7} and CBP_{CFID}) affected acetylation-dependent regulation of CBP condensates. Due to their inability to form visible condensates, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were not affected by AL mutations (Figure S4G-I). However, when expressed on a $CBP_{\Delta IDR6}$-GFP background, no changes in the proportion of punctate nuclei resulting from the AL mutants were observed (Figure 4I), and consistent with the behaviour of $CBP_{\Delta IDR6}$-GFP single mutants, both $CBP_{\Delta IDR6/\Delta AL}$-GFP and $CBP_{\Delta IDR6/KTG}$-GFP formed more condensates per nucleus (Figure 4J) with a higher IID than CBP_{wt}-GFP (Figure 4L), although we note that IID was reduced compared to $CBP_{\Delta IDR6}$-GFP single mutants.

Together, these results highlight the diverse effects of acetylation on CBP condensates. Direct acetylation of CBP can negatively regulate condensates, while indirect acetylation can positively regulate condensates, potentially by providing binding sites for CBP bromodomains. However these results clearly demonstrate the importance of CBP IDRs, and CBP_{IDR6} in particular, for shaping how CBP condensates respond to acetylation.

IDRs regulate CBP-dependent acetylation

Next, we asked whether CBP-IDRs could affect CBP-dependent histone acetylation. To better control expression of CBP-GFP, we made clonal HEK293T lines stably expressing doxycycline inducible CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and CBP_{IDR7}-GFP. The cells showed a clear increase in GFP signal in response to increasing concentrations of doxycycline (Figure S4J-K). For all subsequent experiments we chose doxycycline concentrations that maintained overall CBP-GFP expression at near-endogenous levels (Figure S4K).

Following induction, CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP formed visible condensates, while $CBP_{\Delta IDR7}$-GFP displayed a more diffuse signal (Figure 4M). The overall proportion of punctate nuclei (Figure S4L) and the

351 number of condensates per nuclei (Figure 4N) was lower for inducible CBP-GFP than when transiently
352 expressed (Figure S4M). $CBP_{\Delta IDR6}$-GFP condensates (mean IID = 126.0 ± 21.5 s.e.m.) were also larger and
353 more intense than CBP_{wt}-GFP condensates (mean IID = 80.6 ± 18.3 s.e.m.), while $CBP_{\Delta IDR7}$-GFP failed to
354 form quantifiable puncta (Figure 4M).

355 Next, we tested how CBP-IDRs impacted histone acetylation. Induction of $CBP_{\Delta IDR6}$-GFP increased total
356 H3K27ac and $CBP_{\Delta IDR7}$-GFP reduced total H3K27ac (Figure 4N). To test whether HAT activity was affected
357 by $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP, we purified the induced CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP
358 and tested their ability to acetylate recombinant nucleosomes *in vitro*. We expected that dilution of CBP and
359 disruption of chromatin during sample preparation would prevent formation of condensates *in vitro.*
360 CBP_{wt}-GFP had the highest HAT activity, while H3K27ac was reduced for $CBP_{\Delta IDR6}$-GFP and further
361 decreased for $CBP_{\Delta IDR7}$-GFP (Figure 4O). The observed difference between increased H3K27ac in cells
362 expressing $CBP_{\Delta IDR6}$-GFP, and decreased H3K27ac *in vitro*, likely indicates that total histone acetylation levels
363 in cells is a product of both HAT activity and chromatin localisation patterns.

364 CBP IDRs shape how CBP binds to chromatin.

365 Our results suggest that CBP-IDRs outside the HAT domain can influence HAT activity, and highlight a
366 discrepancy between CBP behaviour *in vitro* and levels of histone acetylation in cells. We reasoned that as
367 well as modulating HAT activity, regulation of CBP condensates by CBP-IDRs could change CBP chromatin
368 association, and thus regulate total levels of histone acetylation. We tested this in our inducible CBP-GFP cell
369 lines using chromatin immunoprecipitation sequencing (ChIP-seq) and CUT&RUN.

370 Firstly, we confirmed that induced CBP-GFP constructs could bind to the same locations as endogenous CBP.
371 We compared ChIP-seq profiles for endogenous CBP_{wt} in maternal $HEK293T_{Tet3G}$ cells (α-CBP; $CBP_{wt\text{-}Tet3G}$)
372 and for CBP-GFP (α-GFP; following addition of Doxycycline to inducible cell lines (Figure 4M). Read
373 densities for CBP ChIPseq ($CBP_{wt\text{-}Tet3G}$) and GFP ChIPseq were well correlated across the genome
374 (CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP; 0.855 < Pearson's coefficient < 0.934, Figure S5A; 0.763 <
375 R^2 < 0.883, Figure S5B), confirming that reads aligned to similar locations. We then used MACS2 to highlight
376 CBP-bound sites in pooled replicates (Figure 5A). We identified 14,913 binding sites for $CBP_{wt\text{-}Tet3G}$ using CBP
377 ChIP-seq, and observed more binding sites for both $CBP_{\Delta IDR6}$-GFP (18,357) and $CBP_{\Delta IDR7}$-GFP (17,237) than
378 for CBP_{wt}-GFP (7,661) in GFP ChIP-seq (Figure 5A). CBP_{wt}-GFP and $CBP_{wt\text{-}Tet3G}$ bound to similar regions in
379 the genome; most binding sites were found at promoters (51.5% CBP_{wt}-GFP; 52.1% $CBP_{wt\text{-}Tet3G}$), then introns
380 (14.4% CBP_{wt}-GFP; 18.6% $CBP_{wt\text{-}Tet3G}$) and intergenic sites (13.5% CBP_{wt}-GFP; 15.0% $CBP_{wt\text{-}Tet3G}$); 20.5% of
381 CBP_{wt}-GFP binding sites and 14.3% of $CBP_{wt\text{-}Tet3G}$ binding sites mapped to exons (Figure 5B). This
382 localisation shifted towards introns and intergenic regions for $CBP_{\Delta IDR6}$-GFP (22.6 % introns, 19.9 %
383 intergenic) and $CBP_{\Delta IDR7}$-GFP (27.2 % introns, 21.8% intergenic), (Figure 5B). $CBP_{wt\text{-}Tet3G}$ and CBP-GFP
384 binding sites were closer than expected by chance[55] (Figure S5D), demonstrating spatial correlation between
385 CBP-GFP and CBP_{wt}; all CBP-GFP binding sites showed statistically-significant overlap with $CBP_{wt\text{-}Tet3G}$
386 (Figure 5A and S5D). As expected, overlap between CBP-GFP binding sites was more robust than between

387 CBP_{wt} and CBP-GFP; CBP_{wt}-GFP sites overlapped 77% of $CBP_{\Delta IDR6}$-GFP and 69% of $CBP_{\Delta IDR7}$-GFP.
388 However, only 43% of $CBP_{\Delta IDR7}$-GFP bound-sites overlapped $CBP_{\Delta IDR6}$-GFP (Figure S5D).

389 Having confirmed that CBP-GFP was binding to similar locations as CBP_{wt}, we next asked how CBP-IDRs
390 affected the profiles of CBP-GFP chromatin binding. Compared to CBP_{wt}-GFP at endogenous CBP binding
391 sites, $CBP_{\Delta IDR6}$-GFP read densities were enriched, while $CBP_{\Delta IDR7}$-GFP was depleted (Figure S5E,F). Across
392 all CBP-GFP binding sites, the mean peak width for $CBP_{\Delta IDR6}$-GFP (485±2.5bp) was greater than for
393 CBP_{wt}-GFP (418±3.2bp), and the peak width for $CBP_{\Delta IDR7}$-GFP (357±2.0 bp) was lower than for CBP_{wt}-GFP
394 (Figure 5C). In summary, the behaviour of CBP constructs when bound to chromatin mirrors their behaviour
395 in imaging data: $CBP_{\Delta IDR6}$-GFP bound to more locations with a larger average peak size, while $CBP_{\Delta IDR7}$-GFP
396 bound to more locations but with a smaller average peak size. The results suggest that larger condensates
397 formed by $CBP_{\Delta IDR6}$-GFP increased its contacts with chromatin, while dispersion of $CBP_{\Delta IDR7}$-GFP
398 condensates led to lower levels of CBP binding at a larger number of locations.

399

400 CBP IDRs shape histone acetylation.

401 Next, we used spike-in normalised CUT&RUN to ask whether altered patterns of chromatin localisation
402 affected H3K27ac levels. Genome browser views of CBP-bound regions and corresponding H3K27ac
403 enrichment, such as at the *TSC22D3* gene and a candidate CRE (cCRE) (Figure 5D), highlight binding of
404 CBP-GFP to similar locations as CBP_{wt} and corresponding enrichment for H3K27ac. Comparison of sites
405 enriched for H3K27ac and for CBP_{wt}-GFP across the genome showed that H3K27ac levels were elevated in
406 CBP_{wt}-GFP compared to $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP (Figure 5E, S5G). However, CBP_{wt}-GFP levels
407 were lower than both $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP at the same loci (Figure 5F, S5G). This suggests that
408 despite higher levels of recruitment at H3K27ac enriched sites, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP were less
409 active, and therefore H3K27ac levels were lower, in line with reduced HAT activity we observed for
410 $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP *in vitro* (Figure 4O).

411 To directly compare how CBP-IDRs affect H3K27ac levels, and avoid complexities from sampling
412 CBP-bound and CBP-unbound locations, we next focussed on a subset of 4,765 consensus binding sites
413 present in all CBP-GFP datasets (Figure 5A). Consensus binding sites predominantly overlapped promoter
414 regions (58%) and exons (22%), while introns (11%) and intergenic regions (9%) were less common (Figure
415 5G). CBP-GFP read densities showed similar patterns of enrichment at consensus sites as seen previously at
416 CBP-bound sites: read densities for CBP_{wt}-GFP and $CBP_{\Delta IDR7}$-GFP were similarly enriched, while
417 $CBP_{\Delta IDR6}$-GFP levels were higher (Figure 5H). H3K27ac read densities also showed similar patterns to those
418 seen previously. The highest levels for H3K27ac were seen with CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP, while
419 expression of $CBP_{\Delta IDR7}$-GFP caused lower levels of H3K27ac, again suggesting that $CBP_{\Delta IDR6}$-GFP and
420 $CBP_{\Delta IDR7}$-GFP were less active (Figure 5I-J).

421 We next asked how CBP-IDRs affected H3K27ac at the majority of sites that were not bound by all CBP-GFP
422 constructs. While only 20% of consensus CBP binding sites were found at introns and intronic regions, this
423 balance shifted considerably for non-consensus sites, with CBP_{wt}-GFP (21% introns, 20% intergenic),
424 $CBP_{\Delta IDR6}$-GFP (27% introns, 23% intergenic) and $CBP_{\Delta IDR7}$-GFP (33% intron, 27% intergenic) all binding to
425 introns and intergenic regions more frequently. The increased proportion of $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP
426 binding to these regions suggests that when CBP forms larger condensates due to deletion of CBP_{IDR6}, and the
427 condensates are dispersed following deletion of CBP_{IDR7}, this causes a relocation of CBP binding to introns
428 and intergenic sites.

429 To assess the effect of CBP-GFP on levels of H3K27ac at non-consensus locations, we compared H3K27ac
430 enrichment at non-consensus sites in CBP-GFP cell lines with the background level of H3K27ac at the same
431 locations in $CBP_{wt\text{-}Tet3G}$ cells (Figure 5L-M). Recruitment of CBP_{wt}-GFP caused a modest enrichment of
432 H3K27ac above background. However, $CBP_{\Delta IDR6}$-GFP significantly increased H3K27ac at non-consensus
433 sites, while $CBP_{\Delta IDR7}$-GFP decreased H3K27ac compared to background levels at non-consensus sites, despite
434 increased binding of $CBP_{\Delta IDR7}$-GFP. This was likely indicative of its lower catalytic activity (Figure 4O), while
435 the slight decrease in the activity of $CBP_{\Delta IDR6}$-GFP could be counteracted by its higher level of recruitment.

436 Finally, as CBP is recruited to CREs across the genome, we examined how CBP-IDRs affected H3K27ac level
437 at cCREs defined by ENCODE[56]. Compared to CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP was enriched and $CBP_{\Delta IDR7}$-GFP
438 levels were reduced at cCREs, consistent with CBP enrichment at H3K27ac domains identified in our
439 CUT&RUN data and at CBP binding sites identified by GFP ChIP-seq (Figure S5I). H3K27ac also showed a
440 similar pattern at cCREs, being most enriched for CBP_{wt}-GFP, then $CBP_{\Delta IDR6}$-GFP and diminished by
441 $CBP_{\Delta IDR7}$-GFP (Figure S5J). Taken together, the data highlight that the combined effect of CBP-IDRs on HAT
442 activity and chromatin localisation determine patterns of histone acetylation at CBP binding sites, and that this
443 behaviour is also seen at cCREs across the genome. The increased binding of $CBP_{\Delta IDR6}$-GFP, together with its
444 lower HAT activity, slightly decreased H3K27ac at regulatory regions compared to CBP_{wt}-GFP, while
445 $CBP_{\Delta IDR7}$-GFP reduced both activity and disrupted binding, causing a larger decrease in H3K27ac at regulatory
446 elements.

447

448 CBP-IDRs affect gene expression

449 Our results show that CBP-IDRs regulate the properties of CBP condensates, and thus shape the localisation of
450 CBP to chromatin and the pattern of histone acetylation at regulatory elements across the genome. As CBP
451 binding to enhancers and altered histone acetylation directly regulate gene expression, we therefore asked
452 whether disrupting CBP-IDRs affected patterns of gene expression. We carried out RNAseq from two
453 biological replicates each of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP following addition of doxycycline
454 to induce their expression (Figure 4M). Replicates from the same sample type were well correlated (Figure
455 S6A), and replicates from each sample formed distinct clusters in principle components analysis (PCA)

(Figure S6B). To identify gene expression changes specifically resulting from disruption of CBP-IDRs rather than induced expression of CBP-GFP, we compared gene expression profiles in $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP to CBP_{wt}-GFP. Expression of $CBP_{\Delta IDR6}$-GFP resulted in significant up regulation of 190 genes, and down regulation of 349 genes compared to CBP_{wt}-GFP (Figure 6A & S6C); expression of $CBP_{\Delta IDR7}$-GFP caused upregulation of 182 genes and down regulation of 278 genes compared to CBP_{wt}-GFP (Figure 6B & S6D).

Importantly, genes with altered expression displayed differences in CBP chromatin binding in GFP ChIP-seq data, and were characterised by changes to H3K27ac levels in spike-in normalised CUT&RUN data. For example *HAND1*, which was upregulated in both $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP datasets, showed additional CBP-GFP binding and elevated H3K27ac sites at regulatory elements (Figure 6C). The results highlight how disruption of CBP-IDRs causes off-target accumulation of CBP and alterations to H3K27ac, leading to widespread deregulation of gene expression.

Discussion

Our results provide direct evidence for the importance of CBP-IDRs as regulators of CBP function. We show that different CBP-IDRs can regulate CBP activity in a sequence-dependent manner, by controlling the properties of CBP condensates, altering chromatin localisation and by shaping histone acetylation and gene expression profiles. Importantly, we uncover very different behaviours in different CBP-IDRs, and show how these IDRs can regulate CBP's responses to lysine acetylation.

The ability of condensates to regulate transcription requires balance: too much condensation downregulates transcription, suggesting that condensates must exist within a window of optimal properties to guarantee accurate transcriptional output[11,13,44,45,57]. Here, we highlight how different IDRs within CBP help maintain this balance. Primary drivers of CBP condensate formation are found in its C-terminal IDRs, where CBP_{IDR6} and CBP_{IDR7} make opposing positive and negative contributions to the formation of CBP condensates, that depend on their amino acid sequence in the case of CBP_{IDR6}. The inhibitory effect of CBP_{IDR6} on the condensates driven by CBP_{IDR7} (Figure 7) is similar to regulatory interactions in those formed by Ataxin-2, where condensate forming IDRs - with prion-like properties similar to CBP_{IDR7} - could be quenched by neighbouring IDRs[58]. In CBP, the balance between CBP_{IDR6} and CBP_{IDR7} strongly influences how CBP interacts with chromatin. CBP and p300 are recruited to chromatin by the association of TFs with TF interacting domains (TFIDs) in the C-terminal regions[59]. Here, we show that disruption of CBP_{IDR6} and CBP_{IDR7} can shape how CBP behaves once it has been recruited to chromatin, potentially by changing its interactions with TF-condensates[11]. Our results show that this leads to two key changes: an increase in the number of CBP binding sites on chromatin, and altered profiles of CBP recruitment. Specifically, broader interactions occur when CBP_{IDR6} is deleted, while CBP binding sites decrease in size following CBP_{IDR7} deletion. Importantly, this behaviour directly mirrors that observed in CBP condensates when CBP_{IDR6} and CBP_{IDR7} are deleted.

These alterations in chromatin binding also reshape histone acetylation patterns at regulatory elements, consequently impacting gene expression. Since CBP functions as a crucial scaffolding factor, modifications to its condensate properties and chromatin binding could further disrupt the recruitment and retention of other vital regulatory factors at these sites(Boija et al., 2017; Gomez-Lamarca et al., 2018; Hunt, Boija and Mannervik, 2022).

We have shown that CBP-IDRs determine the sensitivity of CBP condensates to different forms of lysine acetylation. Again, this regulatory ability is focused on the C-terminal IDRs, where disruption of CBP_{IDR6} causes a larger and longer disruption of CBP condensates by A-485, while de-sensitizing CBP condensates to acetylation events on the AL (Figure 4G). The roles of CBP and CBP-dependent acetylation are tightly linked in gene expression: CBP HAT activity is regulated by autoacetylation on the AL[21] while CBP and p300 can acetylate up to one-third of nuclear proteins as well as acetylating histones [9]. Acetylation has been shown to both negatively regulate condensates formed by the HAT domain of p300 *in vitro* [28] and reduce p300 chromatin association [59], while promoting nucleation of condensates by increasing binding sites for proteins containing bromodomains (including CBP and p300 themselves) [11,52]. Our results demonstrate how C-terminal CBP-IDRs facilitate these processes, by changing the sensitivity of CBP condensates to both direct and indirect acetylation events and by modulating CBP HAT activity. The critical role of C-terminal IDRs in these processes aligns with previous observations that chromatin association, particularly through C-terminal domains, is more readily disrupted by reduced acetylation [59].

Surprisingly, as well as CBP's response to acetylation, disrupting IDRs changed CBP's catalytic activity. CBP-HAT activity was eliminated by deletion of CBP_{IDR7} and drastically reduced by deletion of CBP_{IDR6} (Figure 4O). Truncation of p300 in the TAZ2 domain (N-terminal to CBP_{IDR6} and CBP_{IDR7}) has been shown to increase histone acetylation *in vitro* by relieving TAZ2-dependent inhibition of HAT activity[60]. As TAZ2 remained intact in both our $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP mutants, this suggests that CBP_{IDR6} and CBP_{IDR7} may help to relieve inhibitory interactions between TAZ2 and CBP-HAT. The regulation of HAT activity by C-terminal CBP-IDRs suggests that factors interacting with these regions, such as the NCBD which is located between CBP_{IDR6} and CBP_{IDR7}, may also regulate activity by influencing CBP-IDR behaviour. For example, the adenovirus protein E1A can bind to multiple sites in the N- and C-terminal tails of CBP, but is only able to inhibit the HAT activity of CBP when it interacts with a region overlapping CBP_{IDR6}. Mutations in E1A which specifically disrupt its binding to CBP_{IDR6}, but not other sites in CBP[29], result in the loss of this inhibition. Similarly, stimulation of CBP HAT activity was seen with mutants of pCIP that interact with the C-terminus of CBP via the NCBD, but not by CREB that binds to the N-terminal KIX domain of CBP[29].

Our data has broader implications for the deregulation of CBP function in disease. Importantly, many disease variants in CBP associated with diseases such as the developmental disorders Rubinstein Taybi Syndrome (RTS)[61,62] and Menke-Hennekam Syndrome (MHS)[63], and cancers such as B-cell lymphoma[64], map to IDRs characterised here. Disease associated variants in the HAT domain of CBP are understandably the best characterised, as they are more prevalent and lead to easily definable changes in HAT activity[62,64–66]. However, increasing evidence highlights how variants in IDRs can also lead to disease, including by changing patterns of

condensate formation[67]. This study defines clear roles for CBP-IDRs in regulating diverse cellular functions, whichbif disrupted by mutations could contribute to disease etiology. We also highlight significant sequence divergence between functionally relevant IDRs in CBP and their counterparts in p300. Although not tested here, we speculate that these less well-conserved, but functionally important domains, could form the basis for functional specialisation of CBP and p300.

There are a number of limitations to our study. Firstly, as we did not deplete endogenous CBP, we cannot rule out the possibility of complexities arising from mixed populations of mutant and wild-type CBP. Nevertheless, our work combining CBP mutants with endogenously tagged CBP populations suggests that behaviour we attribute to CBP IDRs can also change endogenous protein behaviour. We also note the complexity inherent in the formation of condensates at regulatory elements, which are capable of binding multiple factors including both CBP and p300[10,11]. Indeed many diseases arising from variants in CBP are heterozygous[64], and therefore the ability of IDRs to confer dominant positive/negative effects on endogenous CBP behaviour that we observe here, may also be relevant to these mixed populations. Our initial work used transient co-expression of CBP-GFP to assess condensate behaviour. This system is imperfect, as the formation of condensates is highly sensitive to protein concentration. Nevertheless, we believe these approaches accentuate behaviours that can also occur at near-endogenous expression levels, as demonstrated using an inducible system to regulate expression. We cannot rule out the possibility that the observed differences in the behaviour of CBP arise because of altered interactions with TFs or other factors, or by disruption of normal patterns of post-translational modifications on CBP-IDRs. CBP_{IDR6} and CBP_{IDR7}, studied in depth here, bridge the NCBD of CBP, a TF binding domain that undergoes coupled folding and binding to associate with different TFs[15–19]. Given that both CBP_{IDR6} and CBP_{IDR7} overlap the NCBD, it is possible that disrupted NCBD folding, leading to altered TF binding, explains the changes in condensate behavior observed in our study. However, we note that deletion of a region equivalent to the NCBD in p300 had only mild effects on p300 chromatin association[59].

CBP is fundamentally important for controlling gene expression, integrating diverse environmental signals to generate precise histone acetylation profiles. Our work highlights how CBP-IDRs facilitate this process by regulating CBP condensate properties, thus enabling responses to various environmental stimuli, such as acetylation, and consequently altering chromatin localisation and gene expression patterns.

Acknowledgements

D.B. and AT are Sir Henry Dale Fellows, funded by the Wellcome Trust and the Royal Society (grant numbers 213501/Z/18/Z and 220192/Z/20/Z respectively); K.G., D.B., A.T. and T.D.C. are supported by a BBSRC Pioneer Award (grant number: BB/Y513453/1); K.G. was supported by a studentship from the White Rose BBSRC Doctoral Training Partnership (grant number: BB/M011151/1); T.D.C was funded by (BB/T008032/1); D.B also received funding from The Royal Society (grant number: RSG\R1\180410). M.D. is part of the Sheffield Bioinformatics Core team. At the University of Sheffield, we thank Susan Clark for

support at the Flow Cytometry Core Facility; IT services for access to High Performance Computing and the Wolfson Light Microscopy Facility. The Nikon W1 spinning disk confocal was funded by a BBSRC ALERT2021 award (grant number: BB/V019368/1). This research was funded in whole, or in part, by the Wellcome Trust [213501/Z/18/Z and 220192/Z/20/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Disclosure and competing interests statement

Timothy Craggs is the founder and CEO of Exciting Instruments, a company that develops and sells instrumentation for single-molecule fluorescence experiments, including smFRET.

Data availability

The ChIP-seq data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266. All DNA constructs used are available from Addgene.

Methods

Cell culture

Culture and maintenance

Human Embryonic Kidney cells stably expressing the SV40 large T antigen (HEK293T, ATCC) and GP2-293 (Clontech) cells were cultured in complete DMEM (10% FBS, 1% Penicillin-Streptomycin (Gibco)) and grown in 5% CO_2 at 37°C. HEK293T cells containing stable doxycycline inducible CBP-GFP, were cultured in complete DMEM (10% Tetracycline free FBS, 1% Penicillin-Streptomycin) including 2 mg/ml G418 (Generon) and 50ug/ml hygromycin (ThermoFisher) . All cell lines were screened every three months for mycoplasma infection.

Transient transfections

Transient transfections were done using either Lipofectamine 3000 (ThermoFisher Scientific) or LipoD293 (TEBU-Bio). Manufacturers recommended protocol was followed for 6-well dishes with the following details: 1.5 µg of DNA was transfected per well and 7.5 µl of Lipofectamine 3000 reagent was used per well. For LipoD293 (TEBU-Bio), the manufacturers protocol for 6-well 35-mm dishes was followed using: 1 µg of DNA and 3 µl of LipoD293 per well.

Retroviral transduction

Lentiviral transduction of CBP-GFP was done using the pRetroX™ Tet-On® 3G inducible expression system (Clontech). CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were cloned into the pRetroX-pTRE3G vector for lentiviral transduction. To generate lentiviral particles, 2ug of CBP-pRetroX-pTRE3G constructs or pRetroX-Tet3G were co-transfected with 2ug of the envelope pVSV-G plasmid into GP2-293 cells using lipofectamine 3000 (ThermoFisher). After 48 hours, with 1 media change after 24 hours, the viral supernatant was collected and concentrated using Lenti-X concentrator (Takara) following manufacturers instructions. To generate polyclonal pRetroX-Tet3G HEK293T cells, 8 ug/ml polybrene (Merck Millipore) was added the concentrated virus and used to infect HEK293T cells. Infected cells were selected using G418 (Generon) at a concentration of 8 mg/ml for 1 week. To generate CBP-pRetroX-pTRE3G cells, polyclonal pRetroX-Tet3G HEK293T cells were transduced with viral particles and 8 ug/ml polybrene, then selected using 8 mg/ml of G418 and 125ug/ml hygromycin (ThermoFisher) for 1 week. Clonal cell lines were generated by inducing expression of CBP-GFP with 2000 ng/ml of doxycycline for 48 hours before fluorescence-activated cell sorting (FACS) based on the GFP signal, using a FACSMelody Cell Sorter (BD Biosciences) at the University of Sheffield Flow Cytometry Core Facility. Clonal lines were cultured in the background of 2 mg/ml G418 and 50ug/ml hygromycin (ThermoFisher).

Doxycycline induction of CBP expression:

For doxycycline validation experiments, HEK293T cells were plated in 10 cm dishes (Sarstedt) for western blot or 35-mm glass bottom dishes (Ibidi) for imaging. 24 hours post-plating the media was changed to media containing the predetermined concentration of doxycycline calculated to ensure induction of CBP-GFP but not increase overall CBP expression levels (Figure S5A-B)), 2 mg/ml G418 and 50ug/ml hygromycin.

Molecular Biology and cloning

All PCR reactions were completed using a 50 ul reaction containing: 0.5 ul PFuUltra II fusion HS DNA Polymerase (Agilent), 1 % PFu Ultra reaction buffer, 5 % dimethyl sulfoxide (DMSO, Sigma), 200mM dNTP mix, 0.2 uM forward and reverse primer, and template DNA. DNA was initially denatured at 95ºC with an extension temperature of 68ºC for DNA less than 10 Kb, and 72ºC for DNA greater than 10 Kb, with 35 extension cycles unless otherwise stated. To confirm correct assembly and presence of mutations, all plasmids were checked by Sanger sequencing and most were checked by Whole plasmid sequencing (Eurofins). All Constructs have been made available through Addgene (https://www.addgene.org/).

Full length human CBP (CBP_{wt})

To generate human CBP_{wt} (transcript variant 1, NCBI Reference: NM_004380.3) tagged with a C-terminal GFP for overexpression experiments (CBP_{wt}-GFP), *Homo sapiens* CBP transcript variant 2 (NCBI Reference NM_001079846.1, Sino Biological HG17295-UT) was cloned in-frame into pAcGFP-N1 (Clontech) using NEBuilder HiFi DNA Assembly reaction protocol (NEB). CBP was amplified using primers: *Fwd: Human CBP_fwd, Rev: Human CBP_rev* and pAcGFP-N1 was amplified using primers: *Fwd: pAcGFP-N1_fwd, Rev: pAcGFP-N1_rev*; (For primer sequences see Supplementary table 1). Assembled constructs were checked by Sanger sequencing (Eurofins). To convert transcript variant 2 to transcript variant 1, the missing sequence comprising exon 5 was ordered as a custom-synthesised DNA from IDT and inserted into CBP-pAcGFP-N1 using the NEB HiFi assembly. CBP exon 5 was amplified using primers: *CBP missing Exon_fwd, CBP missing Exon_rev*. CBP-pAcGFP-N1 was amplified using primers: *Fwd: GFP-CBP_fwd, Rev: GFP-CBP_rev*. Fully assembled CBP_{wt}-GFP was sequenced by Sanger sequencing using sequential primers at 900bp intervals and confirmed using Whole Plasmid Sequencing (Eurofins) to check for correct assembly.

CBP-IDRs for optoDroplet

The optoDroplet construct pHR-mCh-Cry2WT was a gift from Clifford Brangwynne (Addgene plasmid #101221)[39]. To target the optoDroplet construct to the nucleus we used custom synthesised DNA (gBlock, IDT) encoding the SV40 nuclear localisation signal (NLS) to generate pHR-mCh-Cry2WT-NLS[36]. The SV40 NLS was amplified by PCR using primers: *Fwd: SV40_NLS_NotIRD_fwd, Rev: SV40_NLS_SbfIRD_rev*, to introduce restriction sites for NotI and SbfI. Following restriction digestion with NotI and SbfI (NEB),

pHR-mCh-Cry2WT and the synthesised NLS were ligated using T4 DNA ligase (NEB) and the construct was checked by Sanger sequencing. CBP-IDRs were cloned into pHR-mCh-Cry2WT-NLS construct using the NEB HiFi DNA assembly protocol. pHR-mCh-Cry2WT-NLS was amplified using primers: *Fwd: pHR-mCh-Cry2WT-NLS_fwd, Rev: pHR-mCh-Cry2WT-NLS_rev*. CBP-IDRs were amplified from CBP_{wt}-GFP using IDR-specific primers (Supplementary table 1). All constructs were checked using Sanger sequencing.

$CBP_{\Delta IDR}$-GFP deletion mutants

To generate $CBP_{\Delta IDR}$ mutants, primers were designed to amplify CBP_{wt}-GFP in reverse directions to delete residues comprising the identified CBP-IDRs (Supplementary table 1). Amplified PCR products were digested using DpnI (NEB), the ends of the PCR fragments were then phosphorylated using T4 polynucleotide kinase (PNK, NEB) before undergoing ligation using T4 DNA ligase (NEB). Constructs were sequenced by Sanger sequencing to check the IDRs had been correctly removed from the construct.

Autoregulatory Loop (AL) mutants

Two mutants of the Autoinhibitory Loop (AL, $CBP_{1558-1607}$) were designed. Firstly, CBP_{KTG} contained the AL with all lysines mutated to glycine (K1564G, K1565G, K1583G, K1586G, K1587G, K1588G, K1591G, K1592G, K1595G, K1597G, K1605G, K1606G, and K1607G). Secondly, in $CBP_{\Delta AL}$ the AL was replaced with a glycine/serine linker (sequence GSAGSAAGSGQF) to maintain correct folding[23]. $CBP_{4322-5818}$ - which are flanked by unique restriction sites for NruI and PmlI - were synthesised as gBlocks (IDT) containing the mutant AL sequences (Supplementary table 1). The synthesised gBlocks and CBP-pAcGFP-N1 constructs (CBP-GFP for single mutants; $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP, and $CBP_{\Delta CFID}$-GFP for double mutants) were digested using NruI and PmlI (NEB). The digested constructs were dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB), before the inserts and vectors were assembled using T4 DNA ligase (NEB). The assembled constructs were checked by Sanger sequencing.

Doxycycline inducible CBP-GFP

CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP, and $CBP_{\Delta CFID}$-GFP were sub-cloned into a modified pRetroX-pTRE3G plasmid (Clontech) containing a hygromycin resistance cassette using BamHI-HF and NotI-HF (NEB) restriction enzymes. The fragments were annealed using T4 DNA ligase (NEB) before sequencing to check for correct insertion and subsequently by whole plasmid sequencing.

$CBP_{\Delta IDR6H2}$-GFP and Cloning $CBP_{\Delta IDR7H1}$-GFP:

To delete predicted a-helices in CBP_{IDR6} and CBP_{IDR7}, DNA sequences were synthesised containing CBP 5332 - 7674 which resides within the N-terminal GFP tag, and spans two unique endogenous restriction sites for MluI-HF and BssHII (GeneArt, ThermoFisher). Synthesized sequences contained deletions of $CBP_{1951-1978}$ ($CBP_{\Delta IDR6H2}$-GFP) and $CBP_{2187-2216}$ ($CBP_{\Delta IDR7H1}$-GFP). CBP_{wt}-GFP and synthesised sequences were digested with MluI-HF and BssHII (NEB) and purified by agarose gel extraction. The digested CBP_{wt}-GFP vector was

dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB) before assembly using T4 DNA ligase (NEB). Correct insertion was checked by Sanger sequencing.

Cloning of gRNA plasmids

Designed gRNAs (Supplementary table 1) were cloned into pSpCas9(BB)-2A-GFP (PX458). PX458 was a gift from Feng Zhang (Addgene plasmid # 48138) [68]. 100 µM forward and reverse oligo pairs were combined with T4 polynucleotide kinase (PNK) (NEB) and 10X T4 ligation buffer (NEB) containing essential ATP. For annealing, reactions were heated to 37ºC for 30 minutes, followed by 95ºC for a further 5 minutes and then cooled to 25ºC at a ramp rate of -5ºC per minute. Annealed gRNAs were then diluted 1:100 in sterile H_2O in preparation for the ligation reaction.
pX458 plasmid was digested with BbsI-HF (NEB) DNA and annealed gRNA oligos were ligated using T4 DNA ligase (NEB). Constructs were validated using Sanger sequencing.

Construction of stably integrated CBP-HaloTag cell lines.

gRNA design

Single guide RNAs (sgRNAs) targeting the C-terminus of CBP were designed using the UCSC CRISPR targets tool (https://genome.ucsc.edu/). After selection in UCSC, crRNAs were further evaluated using the Evaluation tool on the E-CRISP site (http://www.e-crisp.org/E-CRISP/reannotate_crispr. html). Plasmids encoding CBP-targeting sgRNAs were transfected into HEK293T cells using the Neon electroporation system and 100 µl kit (ThermoFisher), at 1100V, 20 ms pulse width and 2 pulses. After 48 hours, genomic DNA was extracted using QuickExtract solution (Lucigen) according to the manufacturer's instructions and PCR of CBP C-terminal region was carried out using primers *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1).

gRNA validation

PCR products amplifying the C-terminal region of CBP in both edited and unedited cells were cleaned up using the Monarch PCR cleanup kit (NEB). DNA concentrations were measured using the Qubit dsDNA BR kit (ThermoFisher) according to the manufacturer's instructions. 10X NEB buffer 2 (NEB) was added to 200 ng DNA and PCR products were denatured to single strands by heating at 95ºC for 5 minutes. DNA strands were then hybridised by decreasing temperature from 95ºC to 85ºC at a ramp rate of -2ºC/second, then from 85ºC to 25ºC at a ramp rate of -0.1ºC per second. T7 Endonuclease 1 (T7E1) enzyme (NEB) was added and incubated at 37ºC for 15 minutes, followed by quenching with 0.25 M EDTA. The total reaction volume was run on a 1% agarose gel. To analyse Sanger sequencing data of CRISPR edited DNA sequences, .ab1 files of CBP C-terminal PCR products were uploaded to the Inference of CRISPR Edits (ICE, Synthego) tool (https://ice.synthego.com/#/, Conant et al. 2022).

Generation of clonal CBP-Halo HEK293T cell lines:

Validated plasmids encoding CBP gRNA and custom synthesised CBP HDR template plasmid (Supplementary Table 1, ThermoFisher) were co-transfected into HEK293T cells using the Neon electroporation system and 100 µl kit (ThermoFisher) at 1100V, 20 ms pulse width and 2 pulses. After 48 hours, single cell sorting was performed using a FACSMelody cell sorter (BD Biosciences) into 96 well plates. Clones were allowed to grow until confluent and then screened PCR genotyping for successful integration of the Halotag using primers *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1) .

qPCR

Total RNA was extracted from WT and CBP-HaloTag HEK293T cells, using TRIzol (ThermoFisher) and treated with Turbo DNase (ThermoFisher) to remove genomic DNA, then purified using phenol/chloroform and ethanol precipitation. Reverse Transcription was performed using the High Capacity cDNA kit (ThermoFisher) and qPCR was carried out using Power SYBR reaction mastermix (ThermoFisher) and a Quantstudio 12K Flex (ThermoFisher). A mean Ct was calculated using each triplicate reaction and normalised to 18S rRNA expression.

Imaging

Live cell imaging of CBP-HaloTag

HEK293T cells containing endogenous CBP-HaloTag were seeded at 400,000 cells/well on 35 mm glass bottom dishes that have been PLL coated (as previously described). After 48 hours, cells were labelled with either HaloTag ligand TMR (Promega) at a concentration of 50 nM for 15 minutes at 37ºC in the dark. Cells were left for a further 2 hours before staining with Hoechst 33342 (Fisher) in DPBS as described in a previous section, before washing in DPBS and storing in imaging media (FluoroBright DMEM (Fisher) supplemented with a final concentration of 20 nM Gibco HEPES (ThermoFisher Scientific)) immediately prior to imaging. Cells were imaged using the Nikon W1 spinning disc confocal microscope equipped with a temperature stage set to 37ºC. Images were taken using the 100X oil immersion lens (NA 1.45), using the 405 nm laser to image Hoechst using 45 % laser power with a 100 ms exposure, TMR was visualised using the 514 nm laser using 60 % laser power with an exposure of 300 ms. Images were taken every 500 ms in the absence of binning. Captured images and videos were processed ImageJ/Fiji [69].

FRAP of CBP-HaloTag

CBP-HaloTag cells were plated on PLL coated 35-mm dishes, and labelled with Janelia Fluor® 549 (JF549, Promega) at a concentration of 50 nM for 15 minutes. Cells were subsequently washed twice in DPBS and placed in imaging media (as above) for a 15 Yeminute recovery. Cells were imaged using a custom built

746 scanning TIRF/FRAP imaging system (Cairn) with Nikon Ti2-E inverted stand, Cairn multiline laserbank,
747 iLas2 unit and Photometrics Prime 95B sCMOS camera. Imaging was performed at 37ºC in a humified
748 enclosure (OKO Lab). Puncta were imaged in near-TIRF using 561 nm laser with 200 ms exposure at 10%
749 laser power, using a Nikon CFI Apochromat 100X oil NA 1.49 objective. ROIs were drawn with a diameter of
750 1.1 µm (10 pixels), with multiple puncta per field of view. FRAP settings are as follows: 10 repetitions using
751 the 561 nm laser line of weight 2 on 3 % laser power. A time course was taken with the following pattern: 15
752 pre-bleach frames at 2 fps; FRAP (performed as described above); 60 recovery frames at 2 fps; 30 frame at 0.5
753 fps. Images were quantified and processed in ImageJ/Fiji [69].
754
755

756 Imaging of transfected CBP-HaloTag with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP

757 Endogenous CBP-HaloTag cells were seeded at 250,000 cells/well on PLL coated glass bottom 35-mm dishes
758 and transfected after 24 hours with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP using LipoD293, as has been
759 previously described. 24 hours after transfection cells were labelled with Janelia Fluor® 646 (JF646, Promega)
760 at a final concentration of 2 nM for 15 minutes before washing in DPBS and imaging in imaging media. 35
761 mm dishes were labelled sequentially, immediately before imaging. 4-colour control images were also taken
762 where the HaloTag had not been labelled, to check for bleed-through of overexpressed constructs. In this case,
763 the media was changed to imaging media and imaged using laser lines 405 nm, 488 nm, 532 nm and 637 nm,
764 with an exposure time of 200 ms using 10 % laser power. A 4-colour control was also taken with
765 CBP-HaloTag labelled with JF646 and no overexpression, for this all laser lines used the same conditions as
766 before except the 637 nm laser had a power of 80 % to visualise the endogenous protein. For colocalization
767 experiments, 2 colour multidimensional images were captured using the 488 nm laser with 10 % laser power
768 and a 200 ms exposure to visualise GFP tagged constructs, and 637 nm laser with 80 % laser power and 200
769 ms exposure to view endogenously tagged HaloTag. Images were processed in ImageJ/Fiji [69].

770 Imaging of doxycycline inducible clones:

771 For imaging cells were seeded at a concentration of 250,00 cells/well on PLL (poly-L-lysine) coated 35-mm
772 dishes and doxycycline induced for 48-hours. Imaging was performed in imaging media using the Nikon W1
773 spinning disc confocal microscope. Laser lines 405 nm and 488 nm were used to image the Hoechst 33342
774 (Fisher) and therefore the nuclei, and GFP expression respectively. The 405 nm laser was set with a 100 ms
775 exposure with 60 % laser power, and the 488 nm laser was set with 300 ms exposure time and 60 % laser
776 power. For quantification Z-stacks were taken to capture the whole volume of the nuclei in 0.5 µm increments,
777 these were further processed in ImageJ/Fiji [69]. For the plus doxycycline conditions, imaging media was
778 supplemented with the appropriate concentration of doxycycline.

optoDroplet

35-mm glass bottomed dishes (Ibidi) were coated in PLL for 1 hour at 37ºC before being washed twice with dH_2O and seeded with HEK293T cells at a concentration of 200,000 cells/well. 24 hours post-plating the cells were transfected using LipoD293™(SignaGen Laboratories), following manufacturer's guidelines for 35 mm dishes. After 24 hours the media was changed into complete DMEM (10% FBS, 1% Pen/Strep). After a further 24 hours, prior to imaging the media was changed into FluoroBright DMEM (Fisher) with a final concentration of 20 mM HEPES (ThermoFisher). Live cell imaging was performed on Nikon A1 confocal using the 60X oil immersion objective (NA 1.4), equipped with a humidified temperature stage set to 37ºC. Two laser wavelengths were used to induce global activation of the optoDroplet system; 488 nm to induce the dimerisation of the Cry2 and 560 nm to image mCherry. An image was captured every 8 seconds, with subsequent activation of the mCherry for the same time. Time courses were taken for each construct, for a period of up to 3 minutes. 3 time courses were taken per construct per biological replicate, with a total of 3 biological replicates being taken.

Fluorescence recovery after photobleaching (FRAP) in transfected cells.

For FRAP experiments, 250,000 cells/well of HEK293T cells were plated in 35 mm dishes and transfected using LipoD293 as was described for the optoDroplet live cell imaging. 24-hours post transfection the media was removed and the cells were washed in DPBS before leaving the cells in imaging media. Imaging was performed using the custom built scanning TIRF/FRAP imaging system (Cairn) described above using a Nikon CFI Apochromat 100X oil NA 1.49 objective. Puncta were visualised in near-TIRF using the 488 nm laser set to 5 % laser power with a 100 ms exposure. A circular ROI with dimensions of 4.4 x 4.4 µm was drawn over the puncta to be photobleached; there may be multiple regions within each field of view. Photobleaching was performed using the 488 nm laser, with 10 % laser power, using 5 repetitions with a thickness of 5. 10 images were taken prior to FRAP, followed by 700 frames being collected after, with a frame rate of ~2.6 fps.

1,6 - Hexanediol treatment:

Cells were plated and transfected as previously described for FRAP in transfected cells. Images were captured using the Nikon W1 spinning disc confocal microscope using the 100X oil immersion objective. Images were taken using the 488 nm laser with a 200 ms exposure time on 20 % laser power. A 10% solution of 1,6 - Hexanediol (Merck) was diluted from a 50% stock in warm imaging media. Time courses were taken of the constructs over a 3 minute period, imaging every 500 ms. Without pausing imaging, after 30 seconds 1 ml of 10% 1,6 - Hexanediol was added dropwise to the 35 mm dish that contained 1 ml of imaging media for a final concentration of 5% 1,6 - Hexanediol. Imaging continued until the end of the time course. Image quantification and processing was performed in ImageJ/Fiji [69].

Immunofluorescence and fixed cell imaging

Immunofluorescence and fixed cell imaging

Circular coverslips (Ibidi) were acid washed in 0.25% acetic acid before washing and storing in 100% ethanol. Coverslips were coated in poly-L lysine (PLL, Merck) for 1 hour at 37ºC before washing twice in sterile dH_2O. HEK293T cells were grown in 6-well plates at a concentration of 200,000 cells/well containing the coverslips for 48 hours before fixing. If the cells were transfected, transfection was performed 24 hours after plating using Lipofectamine 3000 transfection reagent (ThermoFisher) or LipoD293 (TEBU-Bio) following manufacturers instructions. Cells were washed in DPBS twice before fixing in 2% paraformaldehyde (PFA) for 30 minutes at room temperature in the dark before being washed 3 times in DPBS. For immunofluorescence, coverslips were blocked using a blocking buffer (5% Bovine serum albumin (BSA) with 0.2% Triton in PBS) for 1 hour at 37ºC. Primary antibody (for full list of antibodies see Supplementary Table 2) was prepared in a blocking buffer and incubated at 37ºC for 1 hour. After washing 3 times in PBS, secondary antibodies were diluted in the blocking buffer and incubated for 37ºC for 1 hour. After fixing, coverslips were mounted onto slides (VWR, 631-1554) using VECTORshield that contained 4,6-diamidino-2-phenylindole (DAPI) (Vector laboratories, H-2000-10). Fixed slides were stored at 4ºC. Imaging work was performed at the Wolfson Light Microscopy Facility using the Nikon A1 confocal microscope. All fixed cell imaging was performed using the 60X oil immersion objective (NA 1.4). The microscope has 4 laser wavelengths, typically only 3 of the 4 were used (these include, 405 nm for DAPI, 488 nm for GFP, 562 nm for mCherry and 642 nm for Alexa Fluor 647). Z-stacks were taken every 0.5 µm to cover the depth of the nuclei; a minimum of 3 fields of view were imaged per construct over 2 coverslips, per biological replicate.

Image analysis

Integrated Intensity Density

Before calculating the Integrated density, individual nuclei were isolated and using a blinded approach, the number of nuclei that display either puncta or a diffuse phenotype was determined for each of the constructs. From this a percentage of nuclei which show either puncta or diffuse was calculated.

Integrated density, the intensity multiplied by the area, for each condensate produced by a defined construct was calculated in ImageJ/Fiji [69] using a custom macro. A maximum projection was generated from each Z-stack, and the condensates were identified by applying a manual intensity threshold to each image to generate a mask. The intensity threshold was internally consistent within each experimental repeat, where the threshold was chosen to best isolate condensates across the range of different phenotypes observed within the constructs tested. Within the mask, each independent ROI was labelled, and the area, Integrated Density, and circularity was measured for each ROI.

Once these images had been analysed, the percentage of nuclei which contained a ROI which has an intensity higher than that of the set intensity threshold value was further calculated. This value was described as the transfection efficiency, and was generated to identify which nuclei had passed the threshold value. Filters were then applied sequentially to the data: 1) To exclude overexpressed nuclei, a maximum grey value of 10 was set; 2) To remove regions of high intensity that were too small to represent puncta, an area threshold (<3 pixels/0.36um) was applied; 3) To remove regions of overlapping puncta, a circularity threshold (<0.25) was set. All downstream analysis was done using R in Rstudio; significance p values for comparison of CBP-GFP behaviours (number of condensate per nuclei and integrated density) were calculated using a Kruskal-Wallis test. Graphs and statistical testing was performed using R in RStudio [70].

optoDroplet

A custom ImageJ macro was written to identify puncta formed in the mCherry channel. Images were run through an ImageJ plugin called Stack Registration (StackReg, [71]), and a threshold was applied using RenyiEntropy to generate a binary mask that underwent binary processing of closing followed by opening. The parameters were then set in analyse particles, applying a size range of 50 - infinity, to generate ROIs which equated to the outline of the signal region, which was the nuclei. These ROIs were then applied to the original, raw image that underwent a multi-measure to calculate the standard deviation of the pixels within the nuclei over the time course. Regions found in the first frame were used to monitor the change in the standard deviation over the time course. Average Fold Change was then calculated, as well as rate of change over the first 32 seconds of the time course. The statistics for the rate of change were performed in GraphPad prism, where an unpaired t-test was used to determine statistical significance (Prism [72]). Graphs for visualisation were made using R in RStudio [70].

FRAP in CBP-HaloTag

Images for each FRAP time course were loaded into ImageJ/Fiji [69] and processed using StackReg ([71]) to remove drift in the sample. Over each time course, the intensity was calculated for a region of interest around the bleached area, before being corrected for photobleaching over the time course using a reference region within the same nuclei. Relative intensity was calculated by subtracting the intensity at T = 0, where, this was defined as the first frame after photobleaching, and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated in GraphPad Prism 9 [72], using a non-linear regression with settings one phase association with a variance weighting of $1/Y^2$.

FRAP in transfected cells

Time courses from FRAP of overexpressed protein were quantified using ImageJ/Fiji [69].The intensity was calculated over the time course, using a ROI of 0.22 µm, which is smaller than the condensate to only monitor the intensity within the condensate. The condensate was manually tracked throughout the time course, moving the ROI when the condensate moved. The relative fluorescence recovery was calculated by subtracting the

intensity at T = 0, the first frame post bleaching and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated as described for FRAP of CBP-HaloTag.

Western blotting

HEK293T cells were harvested and after washing twice in PBS the pellet was lysed in RIPA buffer (50 mM Tris-HCl pH 8 stored at 4ºC, 100 mM NaCl, 2 mM $MgCl_2$, 1 % Triton X-100, 0.1 % Sodium deoxycholate, 0.1 % Sodium Dodecyl Sulfate (SDS)) supplemented with 1 mM Dithiothreitol (DTT), 1X Halt protease inhibitor cocktail, 10 mM sodium butyrate and 500 U/µl benzonase. 25 - 50 ug of overall protein containing sample buffer and reducing buffer was loaded onto a 3 - 8% tris-acetate gel (Thermo Fisher Scientific) using tris acetate running buffer (Life Technologies). Gel was transferred onto a membrane using trans-blot turbo mini 0.2 µm nitrocellulose transfer packs (Biorad) before blocking and probing for CBP, GFP and GAPDH (for full list of antibodies see Supplementary Table 2). Western blots were imaged using G:box (Syngene) gel imager. Images were processed using ImageJ/Fiji [69] and assembled using Adobe Illustrator.

Histone acetyltransferase assay

Expression of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP was induced in the respective cell lines by 48 hours treatment with previously determined doxycycline concentrations to ensure equal expression levels (Figure 4M & S4K). Cells were harvested immediately after doxycycline treatment for cell lysis. Cells were washed with PBS and resuspended in NET buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 2 mM $ZnCl_2$) supplemented with 0.1% NP40, 10mM sodium butyrate and Halt™ Protease Inhibitor Cocktail. Lysed cells were incubated with benzonase (1kU) at 37oC for 30 mins, followed by centrifugation at 4oC to obtain the cell lysate. To immunoprecipitate CBP-GFP, 500µg - 1mg cell lysate was incubated with either 25µl ChromoTek GFP-Trap® Magnetic Particles M-270 or 50µl Protein G Dynabeads™. After recovery, beads were resuspended in 500µl wash buffer (NET buffer supplemented with 10mM sodium butyrate and 0.05% NP-40). 300µl of the bead suspension was retained for a Western blot to confirm successful immunoprecipitation of the GFP-tagged protein. The remaining 200 µl was used for the in vitro HAT assay.

Beads were resuspended in HAT buffer (250mM Tris-HCl, 25% glycerol, 0.5mM EDTA, 250mM KCl) supplemented with 0.1M DTT, 10mM PMSF, 1M sodium butyrate, 1mM acetyl CoA and 0.55µg/ml recombinant polynucleosomes (H3.1; Active Motif) and the reaction mixture was incubated for 30 mins at 30ºC. After incubation, loading dye and reducing agent were added before boiling the beads. The resulting solution was loaded on a 12% Bis-Tris gel with MES running buffer for a Western blot to visualise abundance of H3K27ac.

Western Blotting

Cells were incubated with doxycycline for 48 hours to induce expression of CBP-GFP before cell lysis. ***The protocol from here is the same as you've already written in the Western blotting section, except I used Bis-Tris gels to look for H3K27ac/H3 in addition to TA gels for CBP.*** 50µg of overall protein was loaded onto either a 3-8% Tris-Acetate gel with Tris running buffer (to visualise CBP) or a 4-12% Bis-Tris gel with MOPS running buffer (to visualise H3K27ac).

Chromatin Immunoprecipitation (ChIP):

For ChIP experiments, Dox inducible cells were plated in 10 cm dishes, 2 plates per construct. Cells were induced with concentrations of Dox that were identified using western blot as inducing CBP-GFP without altering overall CBP expression levels (Figure S5A-B). . After 48h, cells were crosslinked at room temperature for 10 minutes using formaldehyde (1.1 % final). Crosslinking was quenched using 2.5 M Glycine (Melford, 125 mM final) incubating for 5 minutes at room temperature. Cells were harvested by cell scraping in PBS, and pelleted by centrifugation at 200 rcf for 5 minutes. Cell lysis was performed in ChIP lysis buffer 1 (50 mM HEPES-KOH (Sigma), pH 7.5 140 mM NaCl, 1 mM EDTA, 10 % Glycerol, 0.5 % NP40, 0.25 % Triton X - 100, Complete protease inhibitor cocktail (Roche)) by rotating at 4ºC for 5 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes, re-suspended in ChIP Buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA (Sigma), Complete protease inhibitor cocktail) and incubated at room temperature on a rotator for 10 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes and re-suspended in ChIP lysis buffer 3 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1 % Na-deoxycholate, 0.5 % N-lauroylsarcosine (Sigma), Complete protease inhibitor cocktail) for shearing. Chromatin was sheared to 150 - 300 bp fragments, using the Bioruptor Pico (Diagenode), samples were sonicated for 12 cycles of 30 seconds on, 30 seconds off pulse sonication. Lysates were cleared by centrifugation at 20,000xg for 20 min. Concentrations of lysates was determined by BCA assay (ThermoFisher), IP's were done with 900 µg of sheared protein in a total of 300 µl with ChIP IP buffer (10 mM Tris-HCl pH 8.0, 600 mM NaCl, 1 mM EDTA, 3 % Triton X-100, Complete protease inhibitor cocktail), input reactions were made for each sample with 90 µg of protein in 50 µl of ChIP IP buffer. Immunoprecipitations were performed by incubating with 4 µg of antibody overnight with rotation at 4ºC. Immunocomplexes were recovered by adding blocked Protein G Dynabeads and incubated for 90 minutes at 4ºC with rotation. Beads were washed 5 times in ChIP wash buffer (50 mM HEPES-KOH pH 7.5, 500 mM LiCl, 1 mM EDTA, 1 % NP40, 0.7 % Na-deoxycholate, 0.1 % N-lauroylsarcosine) and 1x in ChIP final wash buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl). Sample was eluted by incubation at 65°C for 30 min in ChIP elution buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM EDTA, 1 % SDS) with rotation in a thermomixer at 900 rpm. Cross-linking was reversed by incubation at 65ºC overnight. Immunoprecipitated DNA was treated with RNase A (0.2 mg/ml final), supplemented with 4 mM $CaCl_2$ for 2 hours at 37ºC, then

Proteinase K (0.2 mg/ml final) for 2 hours at 55ºC. DNA was then purified by phenol:chloroform extraction and ethanol precipitation and resuspended in 1x TE buffer.

ChIP-seq

For ChIP-seq, samples were prepared using NEBNext Ultra™II DNA Library Prep Kit for Illumina (NEB, E7645) following manufacturers instructions. Briefly, fragmented DNA underwent 5' phosphorylation and dA-Tailing to prepare the fragment ends for adapter ligation. Sample concentration was checked using Qubit dsDNA High Sensitivity (HS) kit (Thermo Fisher), to determine adapter dilutions. For input samples the adapters were not diluted, however a 1:25 dilution of adapters in 10 mM Tris-HCl pH 7.5, was performed for immunoprecipitation samples. A cleanup of the adapter ligation was performed using AMPure XP beads (Beckman Coulter) for size selection following the 200 bp protocol. PCR amplification was performed with 3 cycles of amplification for input samples, and 13 cycles of amplification for immunoprecipitation samples. A cleanup of the PCR reaction was performed using AMPure XP beads (Beckman Coulter), following protocol guidelines. Fragment sizes were checked before sequencing using the TapeStation (Agilent), using High Sensitivity DNA ScreenTape (Agilent).

ChIP-seq Data analysis:

ChIP-seq samples were sequenced using paired-end 150 sequencing on an Illumina NovaSeq 6000 platform (Novogene). Raw data quality was assessed using FastQC v0.11.9[73] and adapters were trimmed using Cutadapt v3.4 [74] implemented in Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned to human genome assembly GRCh38 using Burrows-Wheeler Alignment (BWA) v0.7.17-r1188[75]. For downstream analysis, replicates were merged and duplicate reads removed using Picard v2.27.4 (https://broadinstitute.github.io/picard/), before read quality filtering using SAMTools v1.15.1[76]. BigWig files were generated by normalizing per million reads mapped (RPM) and converting to bedgraph format with Bedtools v2.30.0[77], then to BigWig format using UCSC utilities (bedGraphToBigWig v377)[78]. ChIP-seq peaks were called on filtered data using MACS2 v2.2.7.1[79] using the BAMPE option and a broad peak cutoff of 0.001. Called peaks were assigned to annotated genome regions using Homer software (http://homer.ucsd.edu/homer/ngs/annotation.html). All initial processing, alignment and peak calling was completed as part of the Nextflow ChIP-seq pipeline v2.00[80] on the University of Sheffield High Performance Computing (HPC) cluster. Intersection of ChIP-seq peaks, calculation of relative distance plots[55] and Jaccard statistics[55] were done using Bedtools v2.31.0. Control peaks corresponding to each sample were calculated using bedtools shuffle, over a window restricted to 40kb upstream of transcription start sites on the same chromosome. Subsequent analysis was done using a local installation of DeepTools v3.5.5 [81]. Read density coverage was calculated using multiBigwigSummary in 1kb bins over the entire genome; Correlation coefficients between samples were calculated using plotCorrelation; Heatmaps and metaplots displaying signals over a 5kb window aligned to peak centres were generated using computeMatrix and plotHeatmap; Read density enrichment over defined peak regions were calculated using plotEnrichment. For enrichment of

992 H3K27ac over CBP peaks, we used publicly available data: H3K27ac ChIP-seq in wt HEK293T cells (GEO:
993 GSM1249889)(Liu et al. 2013). All downstream analysis was done using R in Rstudio; significance p values
994 for tag enrichment were calculated using a two-sided Mann-Whitney U-test in R[70]. The ChIP-seq data from this
995 publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the
996 identifier GSE268266.
997
998

999 Computational analysis

1000 Identification of CBP-IDRs

1001 To predict disordered regions in CBP, we used DISOPRED3 [82] and PONDR[37] disorder prediction software.
1002 Tested CBP-IDRs were determined using PONDR VSL2 predictor [37], where stretches of amino acids that have
1003 a prediction of disorder above 50% were classed as intrinsically disordered regions [36]. A mutant was also
1004 generated spanning the region where CBP binds to FUS [38] we termed the CBP-FUS Interaction Domain
1005 (CFID).

1006 Analysis of CBP sequence properties

1007 Analysis of sequence properties was carried out using a local installation of localCIDER[43]. Linear composition
1008 plots for full length CBP_{wt} and CBP_{CFID} were calculated using a blob length window of 200bp and 50bp
1009 respectively. PLAAC analysis was carried out using the web-based PLAAC server[50].

1010 AlphaFold structure prediction (version 2.3.2)

1011 All structure predictions were performed using AlphaFold version 2.3.2[35] implemented through a Colab
1012 notebook. We used the monomer model with relaxation and specified five recycles. AlphaFold produces a
1013 per-residue confidence metric called predicted local distance difference test (pLDDT) on a scale from 0 to 100.
1014 pLDDT estimates how well the prediction would agree with an experimental structure; a pLDDT >70 is
1015 considered to correspond to a generally correct backbone prediction[83]. To assess the confidence of predictions
1016 carried out here, we provided a duplicate view of our model coloured by pLDDT score (Figure S4F). All
1017 visualisation was done using UCSF ChimeraX[84].
1018

1019 Data availability

1020 The ChIP-seq data from this publication have been deposited to the GEO database
1021 (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266. All DNA constructs used are
1022 available from Addgene.

1023

Figure 1

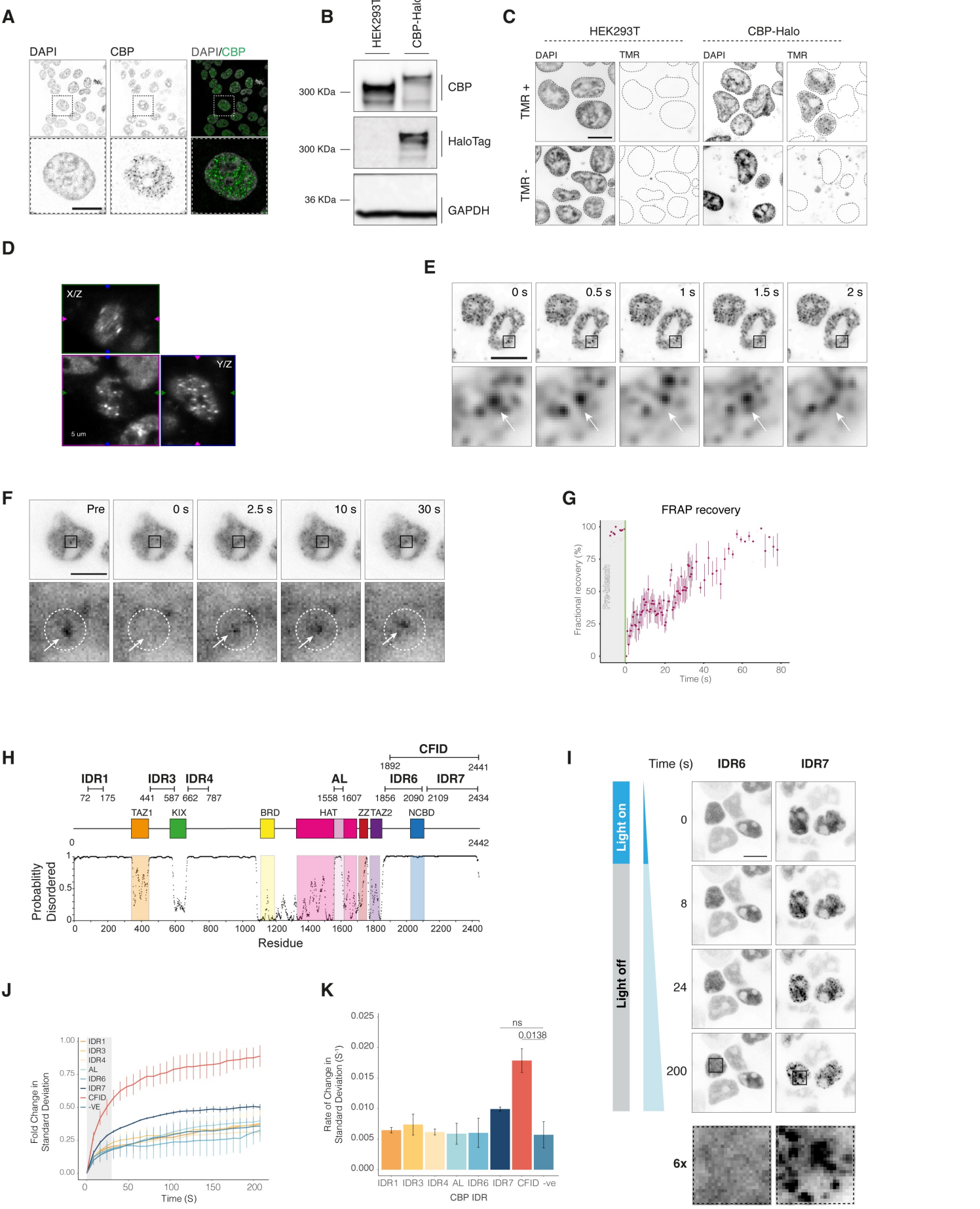

A
DAPI
CBP
DAPI/CBP
B
HEK293T
CBP-Halo
300 KDa
CBP
300 KDa
HaloTag
36 KDa
GAPDH
C
HEK293T
CBP-Halo
DAPI
TMR
DAPI
TMR
TMR +
TMR -
D
X/Z
Y/Z
5 um
E
0 s
0.5 s
1 s
1.5 s
2 s
F
Pre
0 s
2.5 s
10 s
30 s
G
FRAP recovery
Fractional recovery (%)
Pre-bleach
Time (s)
H
CFID
1892
2441
IDR1
72
175
IDR3
441
587
IDR4
662
787
AL
1558
1607
IDR6
1856
2090
IDR7
2109
2434
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
0
2442
Probability Disordered
Residue
I
Time (s)
IDR6
IDR7
Light on
Light off
0
8
24
200
6x
J
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-VE
Fold Change in Standard Deviation
Time (S)
K
ns
0.0138
Rate of Change in Standard Deviation (S-1)
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-ve
CBP IDR

Figure 1: CBP forms fluid-like puncta in the nucleus.

A) Immunofluorescence of CBP in HEK293T cells; Scale bars: 10 µm.

B) Western blot for CBP and HaloTag in HEK293T cells and CBP-Halo HEK293T cells.

C) Confocal microscopy in fixed HEK293T and CBP-Halo HEK293T cells using the fluorescent Halo ligand TMR (CBP-Halo$_{TMR}$); Scale bars: 10 µm.

D) Lattice light sheet microcopy of endogenous CBP-Halo$_{TMR}$ condensates; Single plane images of endogenous CBP-Halo including merged image of Z plane, X/Z and Y/Z planes. Scale bar: 5 µm.

E) Still images from Supplementary Movie 1 showing the dynamic behaviour of CBP-Halo condensates. Boxes indicate regions enlarged in the lower panel; arrows highlight representative dynamic condensates; Scale bars: 10 µm.

F-G) FRAP of CBP-Halo condensates. F) Still images of live cell confocal microscopy showing FRAP of endogenous CBP-Halo$_{JF549}$. Scale bars: 10 µm; G) FRAP curves for endogenous CBP-Halo$_{JF549}$. Data was collected for 32 individual nuclei, and FRAP curves were calculated for 7 condensates where the bleached area did not diffuse away from the focal plane during the recovery period. Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movie 2

H) Domain map and predicted intrinsic disorder (DISOPRED3)[82] of full length CBP$_{wt}$. Positions of CBP-IDRs (CBP$_{IDR1-7}$, CBP$_{AL}$ and CBP$_{CFID}$) with their domain boundaries, and structured CBP domains are highlighted.

I) Live-cell confocal microscopy images showing the different behaviours of CBP$_{IDR6}$ and CBP$_{IDR7}$ in the OptoDroplet assay. Cells were imaged for 200s following exposure to blue light. Dashed box highlights the enlarged region. Scale bars: 10 µm

J) Fold change in the calculated standard deviation of the grey values for defined individual nuclei following blue light stimulation. Diffuse signal should have a low standard deviation, while strong changes in intensity due to condensate formation should have a high standard deviation. Grey shading highlights the 30 second window used to calculate the initial rate of change. Bars represent mean ± s.e.m, n=3.

K) Initial rate of change over the first 30 seconds following blue light exposure. AL, Autoregulatory Loop; -ve, Opto-NLS negative control. Bars represent mean ± s.e.m; significant p values are reported (ns, not significant), n=3.

Figure 2

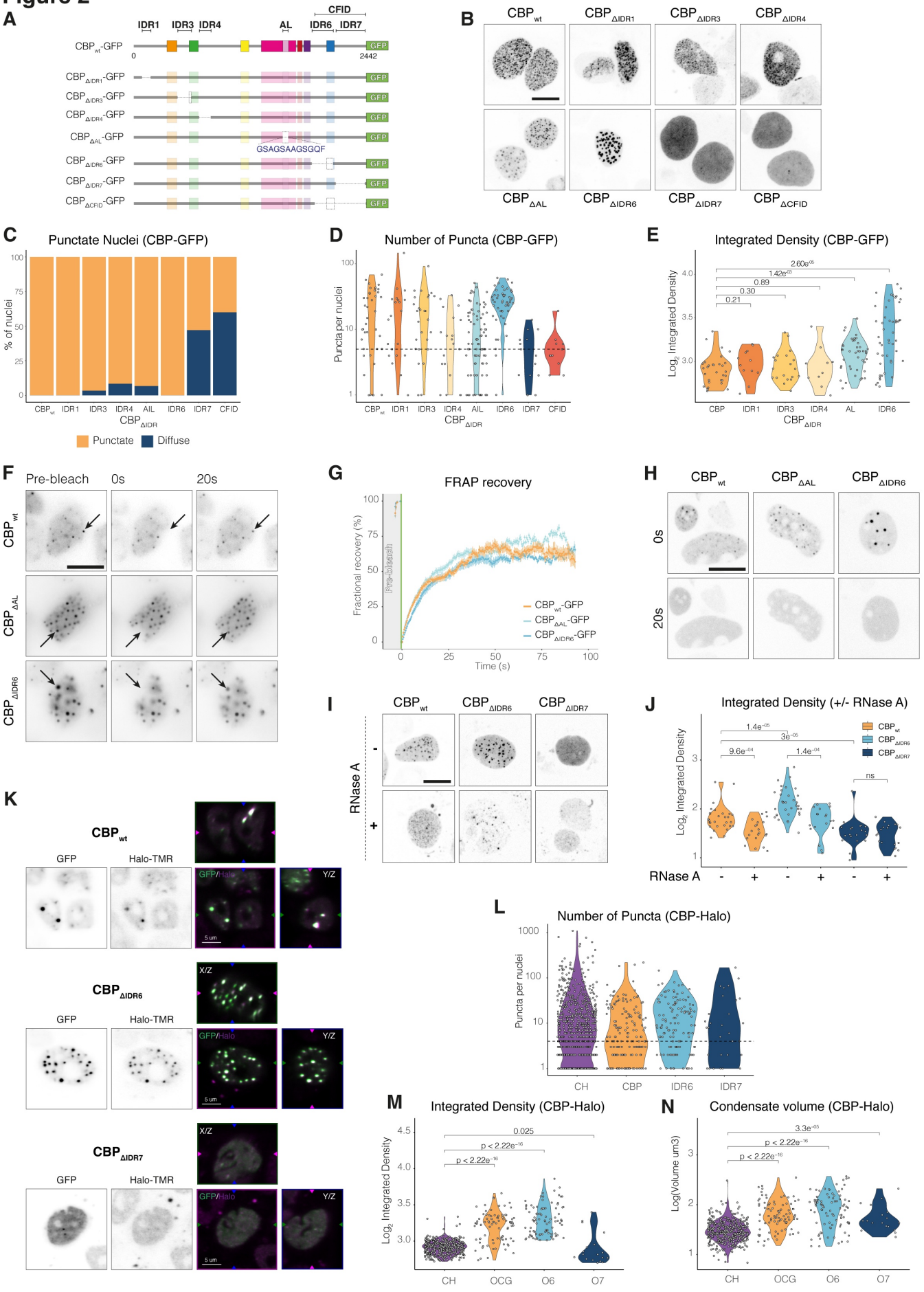

A
CFID
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CBP wt-GFP
0
2442
GFP
CBP ΔIDR1-GFP
CBP ΔIDR3-GFP
CBP ΔIDR4-GFP
CBP ΔAL-GFP
GSAGSAAGSGQF
CBP ΔIDR6-GFP
CBP ΔIDR7-GFP
CBP ΔCFID-GFP
B
CBP wt
CBP ΔIDR1
CBP ΔIDR3
CBP ΔIDR4
CBP ΔAL
CBP ΔIDR6
CBP ΔIDR7
CBP ΔCFID
C
Punctate Nuclei (CBP-GFP)
% of nuclei
CBP wt
IDR1
IDR3
IDR4
AIL
IDR6
IDR7
CFID
CBP ΔIDR
Punctate
Diffuse
D
Number of Puncta (CBP-GFP)
Puncta per nuclei
E
Integrated Density (CBP-GFP)
Log2 Integrated Density
2.60e-05
1.42e-03
0.89
0.30
0.21
CBP
F
Pre-bleach
0s
20s
G
FRAP recovery
Fractional recovery (%)
Pre-bleach
Time (s)
CBP wt-GFP
CBP ΔAL-GFP
CBP ΔIDR6-GFP
H
I
RNase A
J
Integrated Density (+/- RNase A)
1.4e-05
3e-06
9.6e-04
1.4e-04
ns
K
GFP
Halo-TMR
GFP/Halo
X/Z
Y/Z
5 um
L
Number of Puncta (CBP-Halo)
CH
CBP
IDR6
IDR7
M
Integrated Density (CBP-Halo)
0.025
p < 2.22e-16
OCG
O6
O7
N
Condensate volume (CBP-Halo)
3.3e-05
Log(Volume um3)

Figure 2: CBP-IDRs affect the behaviour of full length CBP_{wt}

A) Schematic of $CBP_{\Delta IDR}$-GFP constructs. Deleted regions are highlighted with dashed lines. For $CBP_{\Delta AL}$-GFP, the AL is replaced with a flexible 12 residue linker (blue text).

B) Confocal microscopy in fixed transfected HEK293T cells showing the behaviour of CBP-GFP condensates. Scale bars: 10 µm.

C) Percentage of nuclei displaying diffuse or punctate signal. Images were initially selected for successful transfection based on the presence or absence of GFP signal then the data was blinded and scored on whether GFP signal was punctate or diffuse.

D) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates.

E) Integrated intensity density (IID) of condensates formed by $CBP_{\Delta IDR}$-GFP constructs. p values were calculated using a Kruskal-Wallis test.

F) FRAP analysis of individual CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP puncta. F) Confocal microscopy images showing photobleaching and recovery of CBP-GFP signal. Arrow indicates the bleached puncta. Scale bars: 10 µm.

G) Relative fluorescence recovery normalised to pre-bleached intensity for a 90s time period following bleaching. Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movies 6-8.

H) Effect of 1,6 Hexanediol on CBP-GFP puncta. HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP. Images were taken at fixed time points following application of 5% w/v (final concentration) 1,6 Hexanediol. Scale bars: 10 µm. Still images from Supplementary Movies 3-5.

I-J) RNAse treatment disrupts CBP condensates. Live cell images of transfected HEK293T cells with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. Cells were treated with 100 ug/ml RNase A for 20 minutes, images were taken pre- and post-treatment. Scale bars: 5 µm

K) Lattice light sheet microscopy showing dominant phase separation phenotype on endogenous CBP condensates (JF-646) when HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. Single plane images of endogenous CBP-Halo, and transfected protein including merged image of Z plane, X/Z and Y/Z planes. Scale bar: 5 µm.

Figure 3

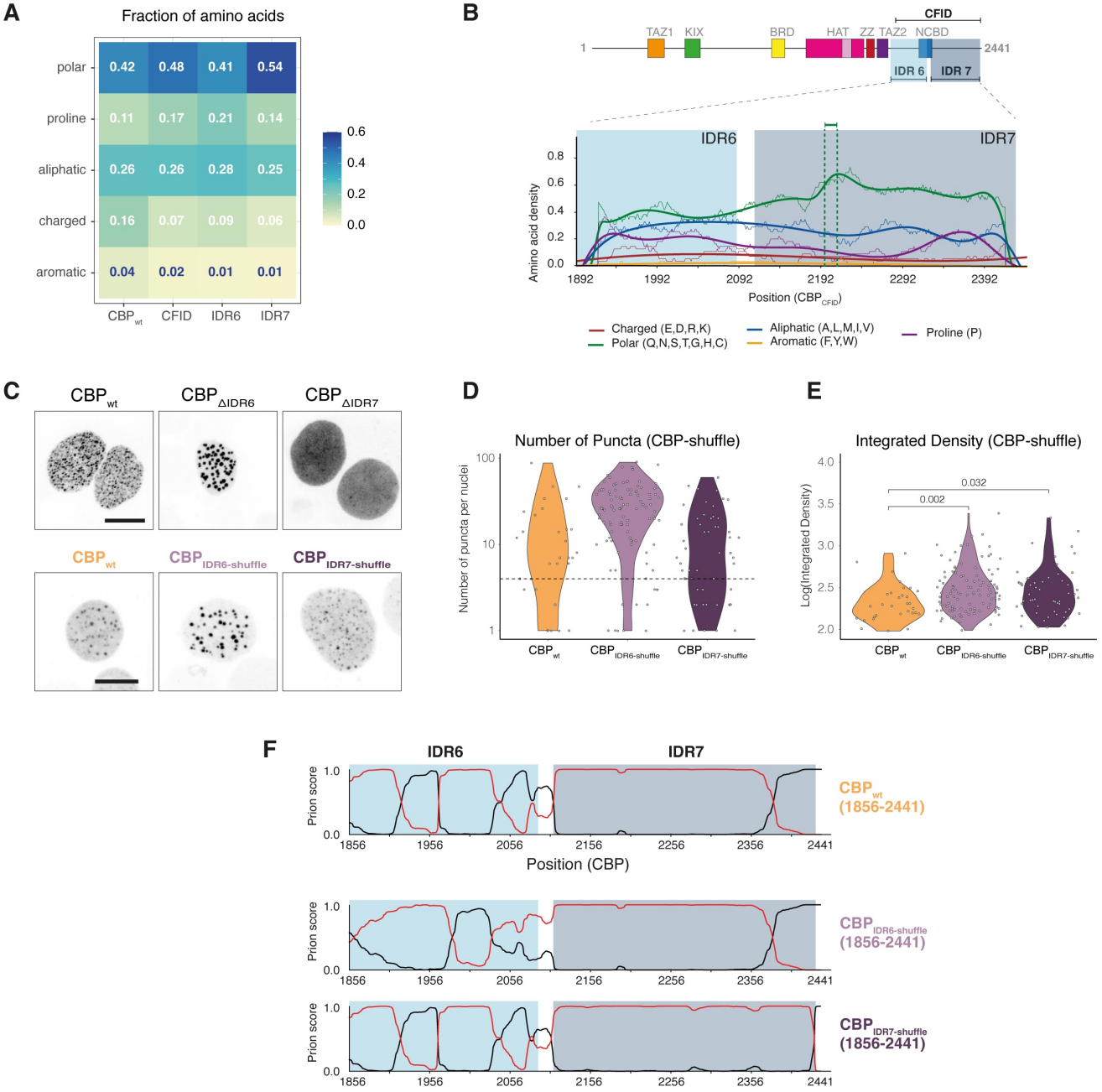

A
Fraction of amino acids
polar
proline
aliphatic
charged
aromatic
0.42 0.48 0.41 0.54
0.11 0.17 0.21 0.14
0.26 0.26 0.28 0.25
0.16 0.07 0.09 0.06
0.04 0.02 0.01 0.01
CBP wt
CFID
IDR6
IDR7
B
CFID
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
1
2441
IDR 6
IDR 7
IDR6
IDR7
Amino acid density
Position (CBP CFID)
Charged (E,D,R,K)
Polar (Q,N,S,T,G,H,C)
Aliphatic (A,L,M,I,V)
Aromatic (F,Y,W)
Proline (P)
C
CBP wt
CBP ΔIDR6
CBP ΔIDR7
CBP wt
CBP IDR6-shuffle
CBP IDR7-shuffle
D
Number of Puncta (CBP-shuffle)
Number of puncta per nuclei
CBP wt
CBP IDR6-shuffle
CBP IDR7-shuffle
E
Integrated Density (CBP-shuffle)
Log(Integrated Density)
0.032
0.002
CBP wt
CBP IDR6-shuffle
CBP IDR7-shuffle
F
IDR6
IDR7
Prion score
Position (CBP)
CBP wt (1856-2441)
CBP IDR6-shuffle (1856-2441)
CBP IDR7-shuffle (1856-2441)

Figure 3: Sequence properties of CBP-IDRs.

A) Sequence properties of CBP_{wt}, CBP_{CFID}, CBP_{IDR6} and CBP_{IDR7}. Amino acids are grouped by property: Polar (Q,N,S,T,G,H,C); Proline (P); Aliphatic (A,L,M,I,V); Charged (E,D,R,K); Aromatic (F,Y,W) based on localCIDER defaults[43]. Numbers indicate the fraction of amino acids in each domain with the grouped property.

B) Linear amino acid composition of CBP_{CFID} calculated using localCIDER[43]. The relative positions of CBP_{IDR6} (light blue) and CBP_{IDR7} (Grey) are highlighted, along a schematic of their positions in the context of CBP_{wt} and CBP_{CFID}; The position of the 18 residue polyQ expansion is highlighted by the green bar and dashed lines. Amino acids are grouped by property: Charged (magenta); Polar (green); Aliphatic (blue); Aromatic (yellow); Proline (purple); Thin lines give the per-residue composition; Sliding window used to calculate local composition=50 residues; Stepsize between windows = 1 residue.

C-E) Effect of Shuffled IDRs on CBP condensates. C) Fixed cell confocal microscopy comparing CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP to their shuffled counterpart. Scale bars: 10 µm; D) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates; E) Integrated intensity density (IID) of condensates. p values were calculated using a Kruskal-Wallis test.

F) PLAAC analysis[49,50] of CBP's C-terminal domains (residues 1856-2441), showing CBP_{wt}, $CBP_{IDR6\text{-}shuffle}$ and $CBP_{IDR7\text{-}shuffle}$. a. The relative positions of CBP_{IDR6} and CBP_{IDR7} are highlighted.

Figure 4

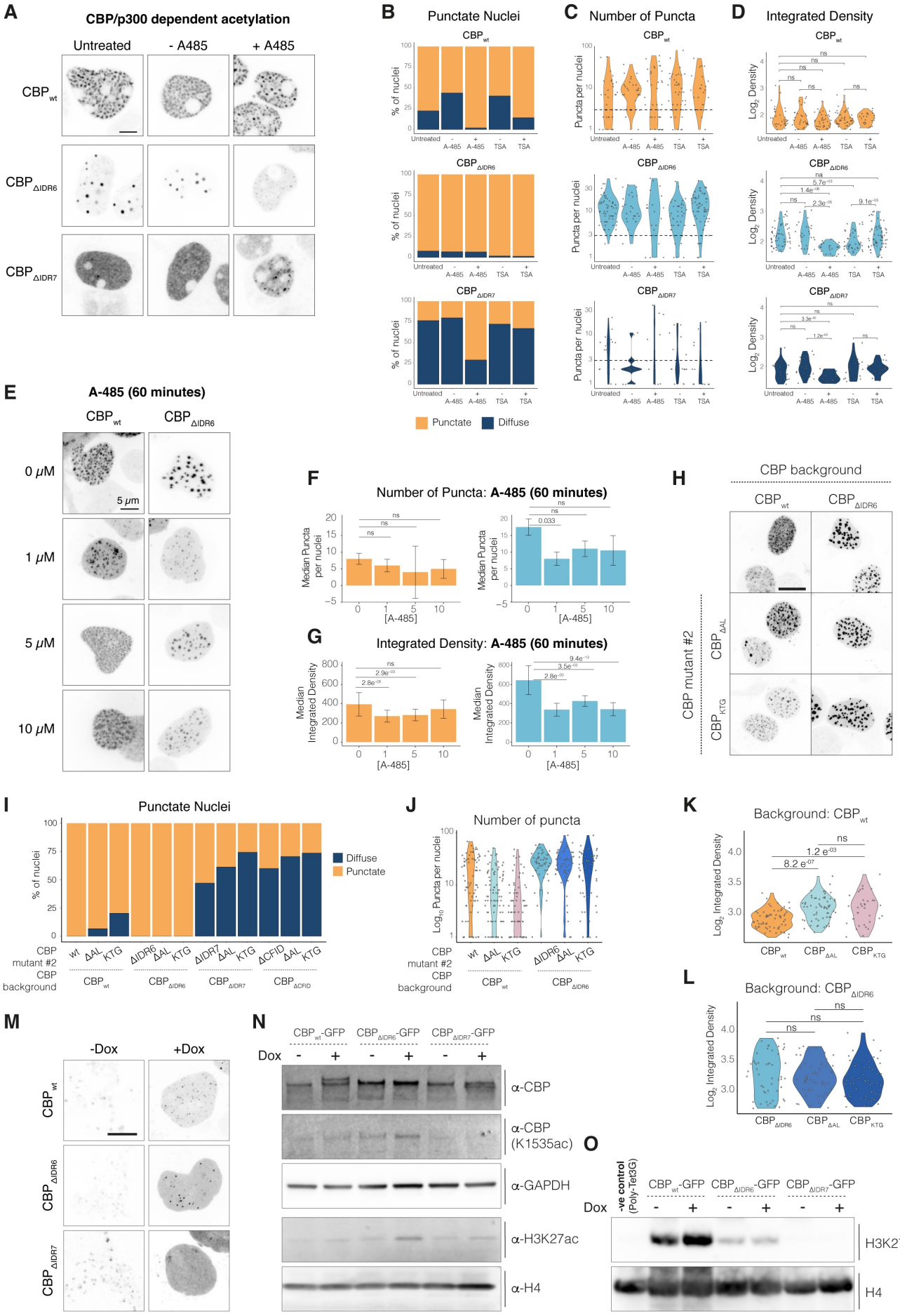

A
CBP/p300 dependent acetylation
Untreated
- A485
+ A485
CBP wt
CBP ΔIDR6
CBP ΔIDR7
B
Punctate Nuclei
% of nuclei
Untreated
A-485
TSA
Punctate
Diffuse
C
Number of Puncta
Puncta per nuclei
D
Integrated Density
Log2 Density
E
A-485 (60 minutes)
0 μM
1 μM
5 μM
10 μM
5 μm
F
Number of Puncta: A-485 (60 minutes)
Median Puncta per nuclei
[A-485]
G
Integrated Density: A-485 (60 minutes)
Median Integrated Density
H
CBP background
CBP mutant #2
CBP ΔAL
CBP KTG
I
Punctate Nuclei
% of nuclei
Diffuse
Punctate
CBP mutant #2
CBP background
J
Number of puncta
Log10 Puncta per nuclei
K
Background: CBP wt
Log2 Integrated Density
L
Background: CBP ΔIDR6
M
-Dox
+Dox
N
CBP wt-GFP
CBP ΔIDR6-GFP
CBP ΔIDR7-GFP
Dox
α-CBP
α-CBP (K1535ac)
α-GAPDH
α-H3K27ac
α-H4
O
-ve control (Poly-Tet3G)
H3K27ac
H4

Figure 4: CBP-IDRs affect response to acetylation

A) Fixed cell confocal microscopy of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP after 48h treatment with 0.5uM A-485, DMSO or no treatment. Scale bars: 5 µm.

B-D) Effect of A-485 or Trichostatin A (TSA) on CBP-GFP condensate behaviour: B)Percentage of nuclei displaying diffuse or punctate signal; C) Number of condensates per nuclei following filtering. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates; D) Integrated intensity density (IID) of condensates formed by $CBP_{\Delta IDR}$-GFP constructs.

E-G) Effect of A-485 concentration on CBP condensate behaviour: E) Fixed cell confocal microscopy of CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP after 60 minutes treatment with stated concentration of A-485 (0um = DMSO). Scale bars: 5 µm; F) Number of condensates per nuclei; G) Integrated intensity density (IID) of condensates.

H-L) CBP-IDRs affect CBP's response to acetylation on the AL: H) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm; I) Percentage of nuclei displaying diffuse or punctate signal; J) Number of condensates per nucleus; K) Integrated intensity density (CBP_{wt}-GFP background); L) Integrated intensity density ($CBP_{\Delta IDR6}$-GFP background).

M) Confocal microscopy showing inducible CBP-GFP expression. Stable inducible CBP-GFP HEK293T cells were treated with Doxycycline (Dox) for 48h. Dox concentrations: CBP_{wt}-GFP, 25ngml-1; $CBP_{\Delta IDR6}$-GFP, 5ngml-1; and CBP_{IDR7}-GFP, 25ngml-1; Scale bars: 10 µm.

N) Western blot of CBP, K1535ac in the AL of CBP/p300 and H3K27ac and following the addition of Doxycycline (Figure 4M & S4K) to induce expression of CBP-GFP. CBP was resolved using 3-8% Tris-Acetate PAGE (loading control: GAPDH, imaged using chemiluminescence) and H3K27ac was resolved using 12% Bis-Tris PAGE (loading control: Histone H4).

O) Immunoprecipitation HAT assay using recombinant di-nucleosome substrate. CBP-GFP expression was induced at equivalent levels using Doxycycline as in (Figure 4M & S4K) and CBP-GFP was immunoprecipitated using GFP-Trap® Magnetic beads. The reaction was resolved by 12% Bis-Tris PAGE and probed for H3K27ac and H4 by western blot.

All p-values for IID were calculated using Kruskal-Wallis tests.

Figure 5

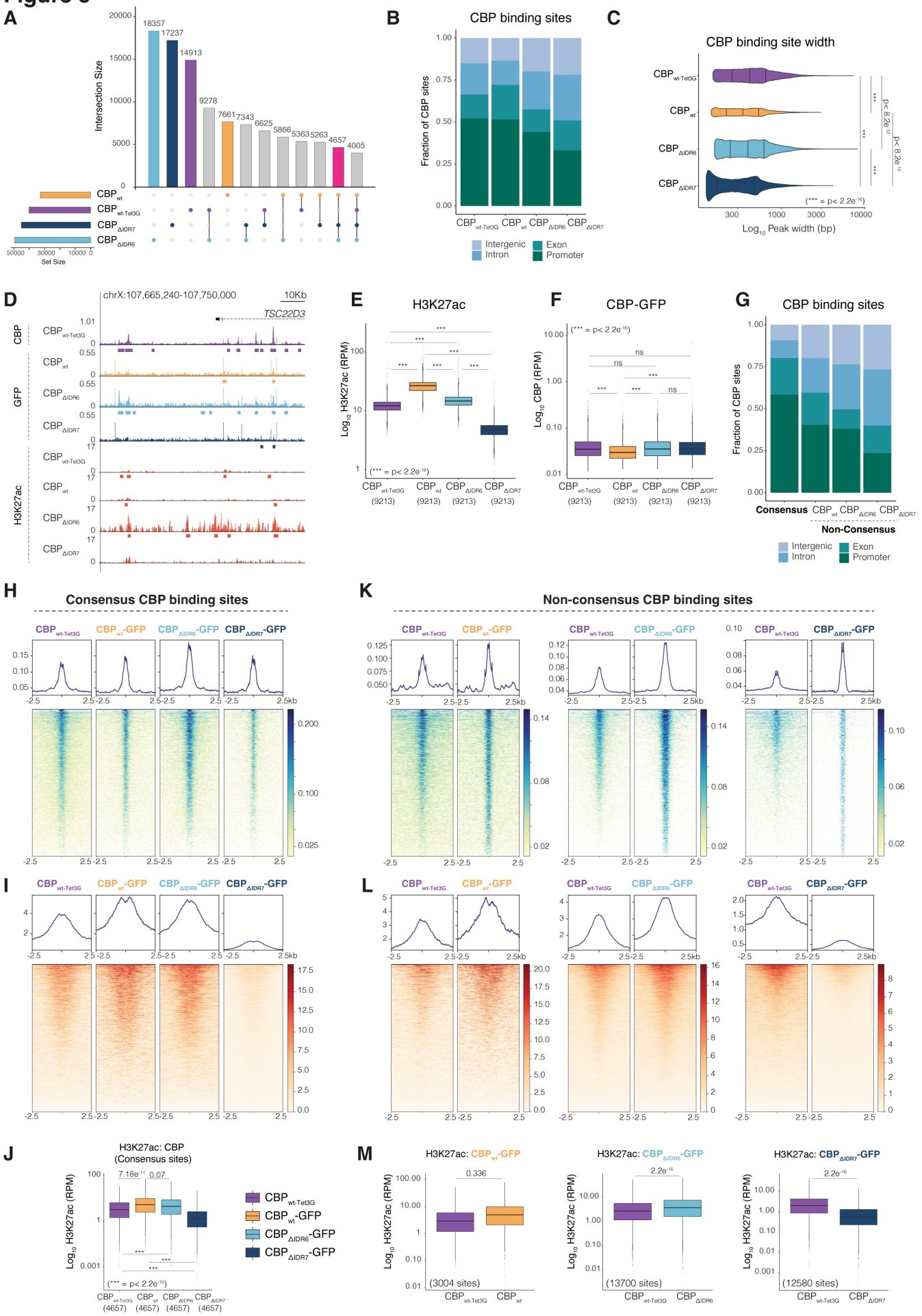

A
Intersection Size
18357
17237
14913
9278
7661
7343
6625
5866
5363
5263
4657
4005
Set Size
B
CBP binding sites
Fraction of CBP sites
Intergenic
Intron
Exon
Promoter
C
CBP binding site width
Log10 Peak width (bp)
(*** = p< 2.2e-16)
D
chrX:107,665,240-107,750,000
10Kb
TSC22D3
CBP
GFP
H3K27ac
E
H3K27ac
Log10 H3K27ac (RPM)
(9213)
F
CBP-GFP
Log10 CBP (RPM)
G
CBP binding sites
Consensus
Non-Consensus
H
Consensus CBP binding sites
K
Non-consensus CBP binding sites
I
L
J
H3K27ac: CBP
(Consensus sites)
(4657)
M
H3K27ac: CBPwt-GFP
H3K27ac: CBPΔIDR6-GFP
H3K27ac: CBPΔIDR7-GFP
(3004 sites)
(13700 sites)
(12580 sites)

Figure 5: CBP-IDRs regulate nuclear chromatin association and H3K27ac.

A) Overlap between CBP binding sites identified from CBP and GFP ChIPseq. $CBP_{wt\text{-}Tet3G}$ (purple), CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). Magenta bar highlights consensus peaks found in every CBP-GFP dataset.

B) Distribution of CBP binding sites by genome region.

C) Peak widths of CBP binding sites. Violin plots show CBP peak widths in $HEK293T_{Tet3G}$ cells (purple), and GFP ChIPseq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). Vertical lines indicate position of 25% quantiles. p values from Mann-Whitney U-test.

D) Genome browser views of CBP and GFP ChIP-seq (top) and H3K27ac CUT&RUN (bottom) at the *TSC22D3* gene. $CBP_{wt\text{-}Tet3G}$ (Purple), CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) CBP_{IDR7}-GFP (dark blue) and H3K27ac (red). Solid bars show binding sites identified using MACS2.

E-F) Read density enrichment at H3K27ac domains. E) H3K27ac; F) CBP-GFP. H3K27ac peaks were randomly downsampled to match the size of the smallest dataset (CBP_{wt}-GFP, 9213); p values from Mann-Whitney U test.

G) Distribution of consensus and non-consensus CBP binding sites by genome region.

H-I) Heatmaps of 4,657 consensus CBP peaks. Heatmaps show ±2.5 Kb window from the centre of peaks. H) CBP ChIP-seq signal; I) H3K27ac CUT&RUN signal.

J) H3K27ac read density enrichment at consensus CBP peaks. p values from Mann-Whitney U test.

K-L) Heatmaps of non-consensus CBP peaks comparing background signal in $CBP_{wt\text{-}Tet3G}$ and enrichment following induction of CBP-GFP. Heatmaps show ±2.5 Kb window from centre of peaks; CBP_{wt}-GFP (3,004 sites, orange), $CBP_{\Delta IDR6}$-GFP (13,700 sites, light blue) CBP_{IDR7}-GFP (12,580 sites, dark blue). K) CBP ChIP-seq signal; L) H3K27ac CUT&RUN signal.

M) Pairwise comparison of background H3K27ac read densities in $CBP_{wt\text{-}Tet3G}$ and CBP-GFP cells at non-consensus CBP peaks. Colours and numbers of peaks are the same as K-L.

Figure 6

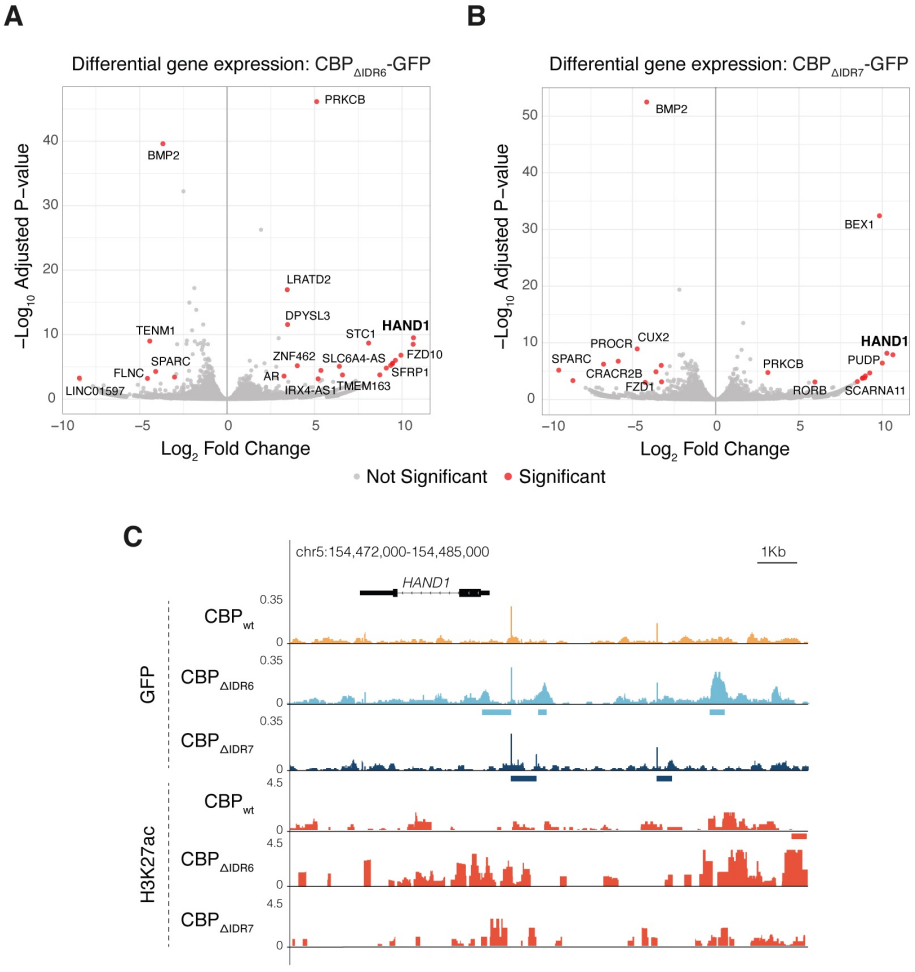

A
Differential gene expression: CBP ΔIDR6-GFP
−Log10 Adjusted P−value
Log2 Fold Change
PRKCB
BMP2
LRATD2
DPYSL3
TENM1
STC1
HAND1
SPARC
ZNF462
SLC6A4-AS
FZD10
FLNC
AR
SFRP1
LINC01597
IRX4-AS1
TMEM163
B
Differential gene expression: CBP ΔIDR7-GFP
BMP2
BEX1
PROCR
CUX2
HAND1
SPARC
PRKCB
PUDP
CRACR2B
FZD1
RORB
SCARNA11
Not Significant
Significant
C
chr5:154,472,000-154,485,000
1Kb
HAND1
GFP
H3K27ac
CBP wt
CBP ΔIDR6
CBP ΔIDR7

Figure 6: CBP-IDRs regulate gene expression profiles

A-B) Differentially expressed genes between A) CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP; and B) CBP_{wt}-GFP and $CBP_{\Delta IDR7}$-GFP. Genes showing significant changes in expression using more stringent significance parameters $p < 0.001$ and $log_2FC > 3$ are highlighted (red points, gene names). HAND1 (bold) is upregulated in both $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP datasets.

C) Genome browser views of GFP ChIP-seq (top) and H3K27ac CUT&RUN (bottom) at the *HAND1* gene. CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) CBP_{IDR7}-GFP (dark blue) and H3K27ac (red). Solid bars show binding sites identified using MACS2.

Figure 7

A

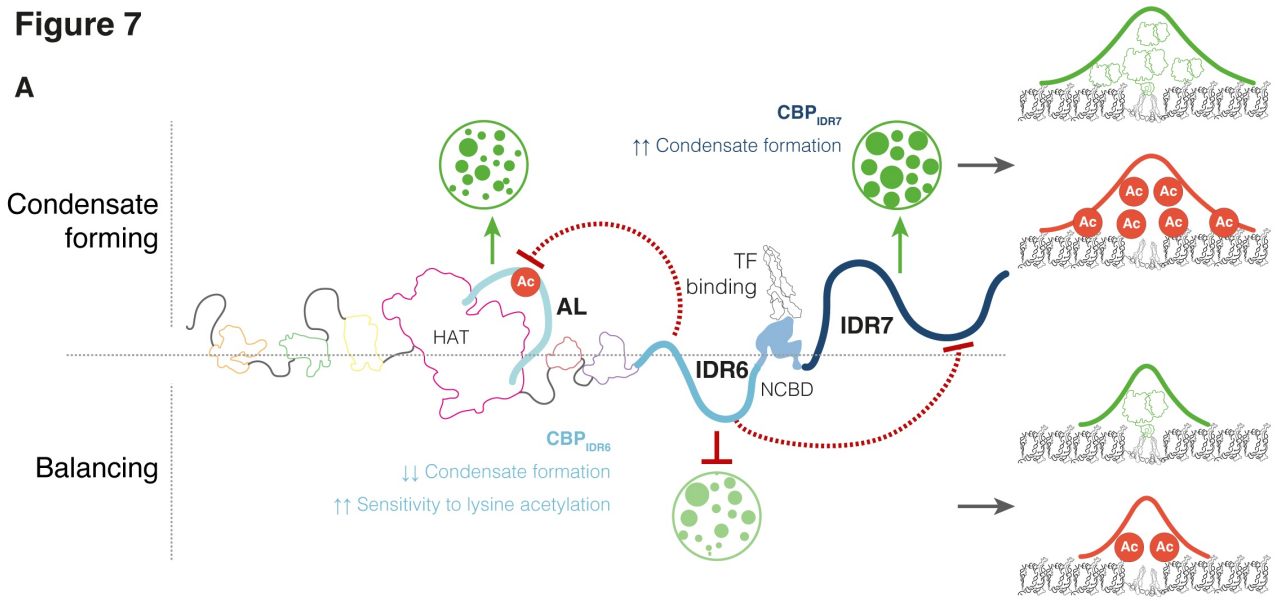
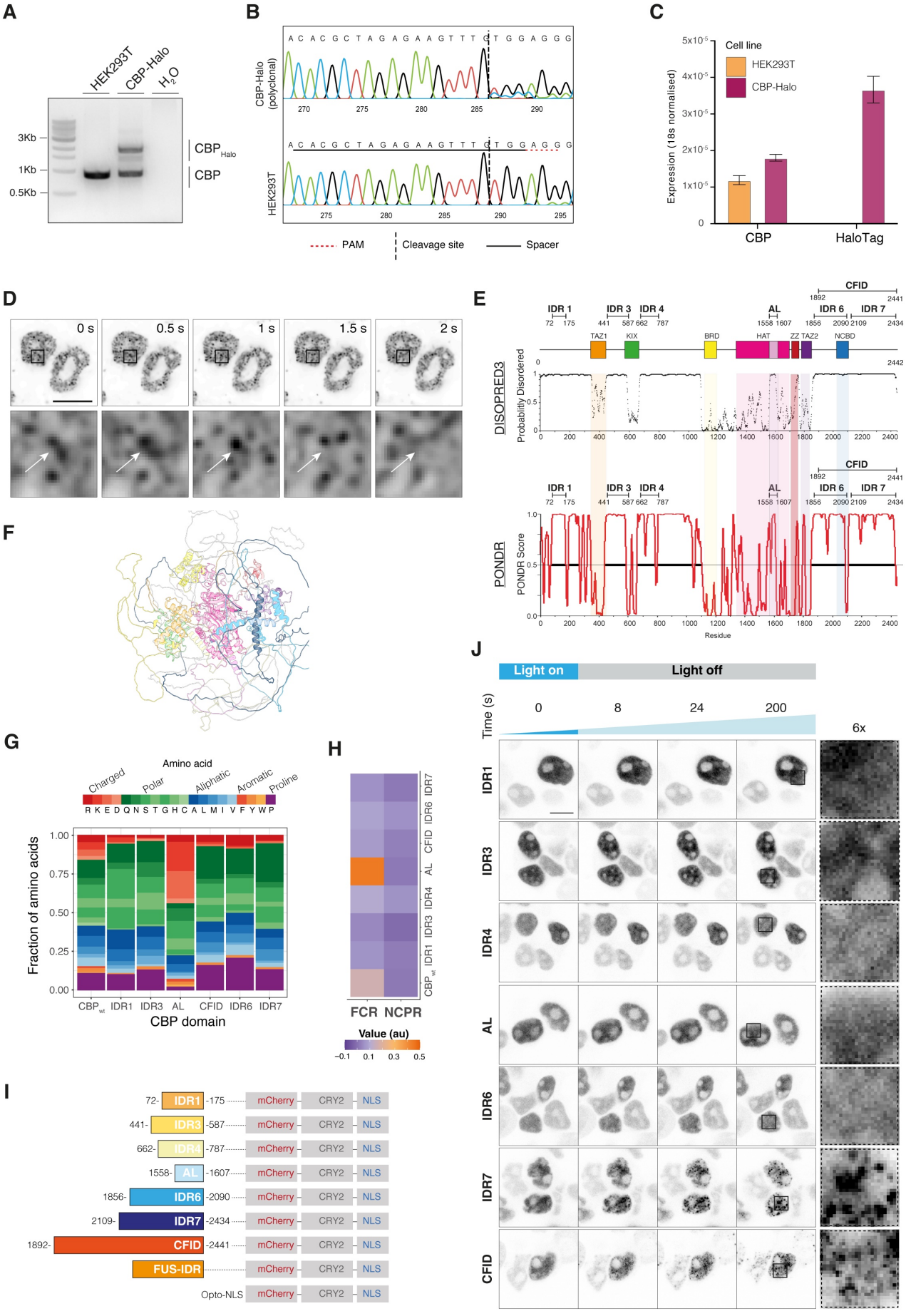

Condensate forming
Balancing
HAT
Ac
AL
TF binding
IDR6
NCBD
IDR7
CBP IDR7
↑↑ Condensate formation
CBP IDR6
↓↓ Condensate formation
↑↑ Sensitivity to lysine acetylation
Ac

Figure 7: Summary of CBP-IDR functions

Behaviours arising from C-terminal CBP-IDRs and their effect on CBP condensates and chromatin localisation.

Supplementary Figure 1



1054 Figure S1

1055 A-B) Analysis of PCR products from edited locus: (A) PCR using primers spanning the 3' end of CBP_{wt}; (B)
1056 Inference of CRISPR Edits (ICE)[85] analysis of Sanger sequencing data from polyclonal CBP-Halo HEK293T
1057 (Top) and HEK293T control cells (bottom).

1058 C) RT-qPCR using primers for CBP or HaloTag in control HEK293T cells (orange) and CBP-Halo HEK293T
1059 cells (magenta). Bars represent mean ± s.e.m, n=3.

1060 D) Still images from Supplementary Movie 1 showing the dynamic behaviour of CBP-Halo condensates -
1061 Alternative representative condensates from (Figure 1E). Boxes highlight regions enlarged below; arrows
1062 highlight representative dynamic condensates; Scale bars: 10 µm.

1063 E) Domain map and predicted intrinsic disorder of full length CBP_{wt} using: DISOPRED3[82], Top; PONDR[37],
1064 bottom. Positions of CBP-IDRs (CBP_{IDR1-7}, CBP_{AL} and CBP_{CFID}) with their domain boundaries predicted using
1065 PONDR, and structured CBP domains are highlighted.

1066 F) AlphaFold v2.3.2[35] prediction for CBP_{wt}. Domains are coloured the same as Figure 1H.

1067 G-H) Sequence properties for CBP_{wt} and CBP-IDRs calculated using localCIDER[43]. G) Amino acid
1068 composition of CBP_{wt} and tested CBP-IDRs; H) Fraction of Charged Residues (FCR, top) and Net Charge per
1069 Residue (NCPR, bottom).

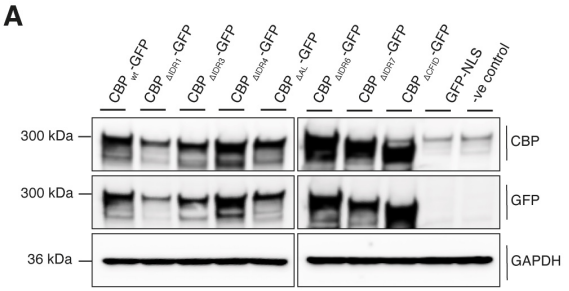
1070 I) CBP-IDR constructs used for OptoDroplet assay.

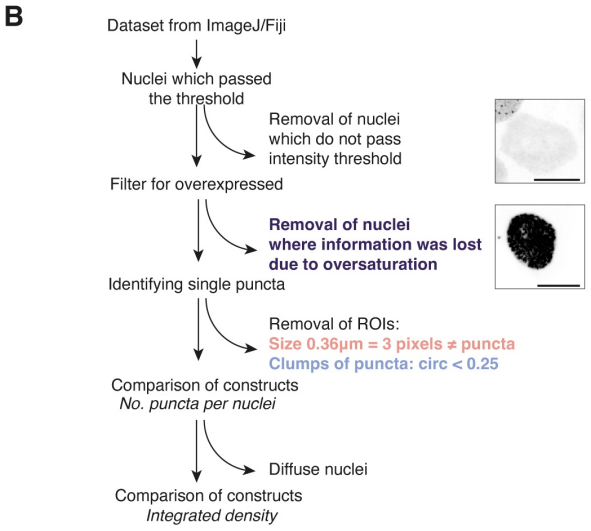
1071 J) Live-cell confocal microscopy images showing behaviour of all tested CBP-IDRs in OptoDroplet assay.
1072 Cells were imaged for 200s following exposure to blue light. Scale bars: 10 µm

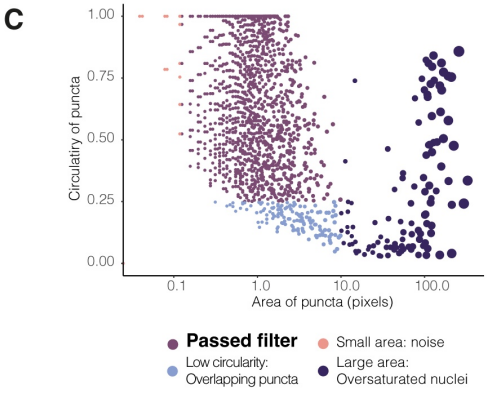
1073

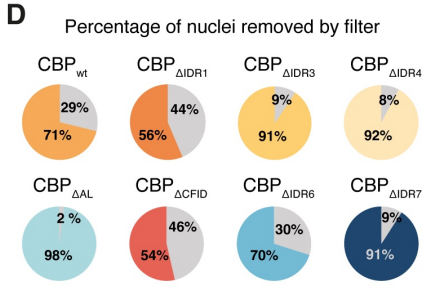
1074

Supplementary Figure 2





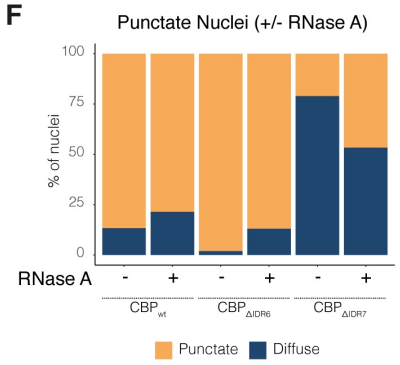


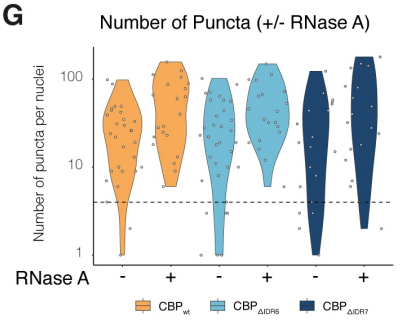


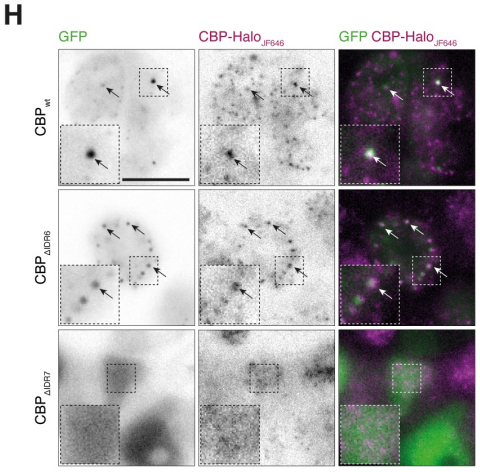
E

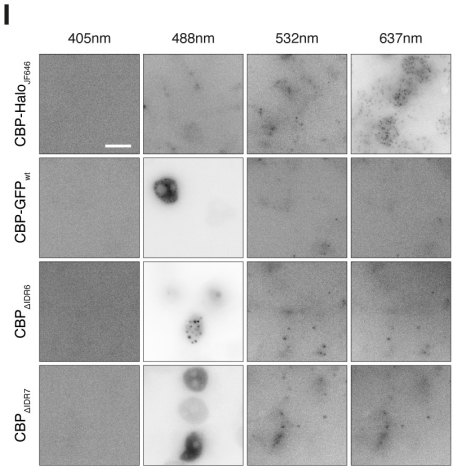
FRAP recovery

CBP-GFP	Recovery rate (s^{-1})	$T_{1/2}$(s)
CBP_{wt}	0.05602	12.37
$CBP_{\Delta AL}$	0.06247	11.09
$CBP_{\Delta IDR6}$	0.06824	10.16









1107 Figure S2

1108 A) Western blot for CBP and GFP shows relative expression levels of transiently transfected CBP-GFP
1109 constructs.

1110 B) Filtering strategy used for selecting transfected nuclei for analysis. Transfected nuclei were selected based
1111 on an intensity threshold [69]. GFP puncta were then filtered by: i) Area - to remove overexpression artefacts (e.g
1112 oversaturated fluorescent nuclei); ii) noise less than 3 pixels in diameter; iii) Circularity - to isolate single from
1113 overlapping puncta. Images show examples of nuclei removed by the filter, scale bars: 10 µm.

1114 C) Summary of filtering results. Highlighted populations are: Nuclei passing filter (Purple); Large area to
1115 remove oversaturated cells/nuclei with no clear puncta (dark blue); low circularity (<0.25) to remove
1116 overlapping puncta (light blue); Puncta < 3 pixels to remove speckle noise (salmon).

1117 D) Percentage of transfected nuclei for each condition remaining following application of the filter.

1118 E) Calculated values for recovery rate over the time course following bleaching and the time to reach 50%
1119 pre-bleach intensity ($T_{1/2}$).

1120 F) Percentage of punctate nuclei following RNAs treatment.

1121 G) Number of CBP condensates per nucleus following RNAse treatment.

1122 H) CBP-IDR mutations have a dominant negative/positve effect on the behaviour of endogenous CBP
1123 condensates. CBP-Halo HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP or $CBP_{\Delta IDR7}$-GFP for
1124 24h. Endogenous CBP-Halo was labelled with Janelia Fluor® 646 Halo ligand ($CBP\text{-}Halo_{JF646}$).

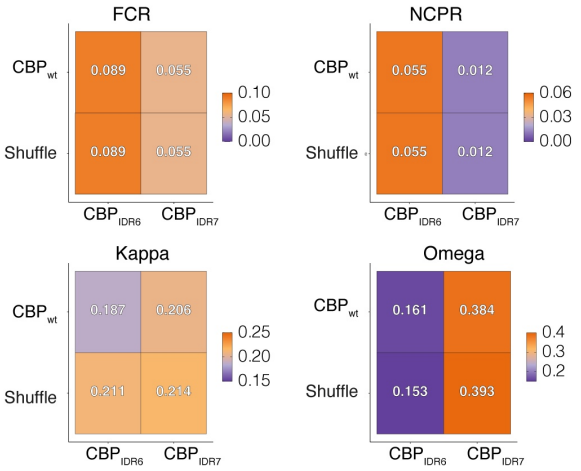
1125 I) Multiple channel control for live cell imaging in Figure S2H. CBP-Halo was labelled with Janelia Fluor®
1126 646 Halo ligand ($CBP\text{-}Halo_{JF646}$).

1127

1128

Supplementary Figure 3

A



B

IDR6 Shuffle (residues 1856-2090)

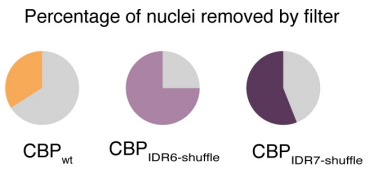
PGLPSPVQNLLPSPSMQSIRAQVATQPQPH-
QVSEGPAAAPRMLNPLQQRGPHPAAMVP-
PVPGSPQPTPGSPQADPRSPLTPPQQMQPTLTL
QQNSGQMQAEVSGPMQAPVTQPPQPIQQMPSRQ
RPRSAATNSFTWGPPVGTIVPAQMQRMMVAPPP
RGSKPAIPSLQERQRPVTRNRPQSSYSSSPVTT
PQSKQANTMLPRNQANMPSIVVVARTLQQPTLQ
SQQAPALGPQPP

C

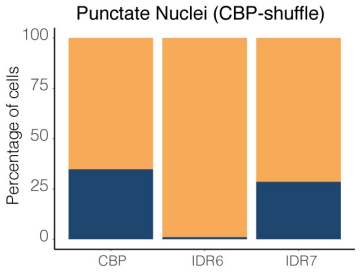
IDR7 Shuffle (residues 2109-2434)

FQPQGDNVQPPQGSLAQQYEQSQLPQAMDMP-
GLMGQSSGPPPTETPGAPPAPQGRDGPAS-
QQMQPLQSARPPAVSQQRMMNTQQMGSQGQQGV
HQLMRGGRQQGQQPGLNQVMAHALPQMHMLQIQ
LLPQSTQQSGMQESQQIGQQPQQMQQHASHKNG
QGQSQSSQNINSSSIRSLLPAPSTQPPISMPGA
AANPPPQLANSLQHISQQTPQQVQYGPRQGAPI
VGPPQMSLQLPQLQAQQLPHGSAHGLLGRHSLT
MQPMGGMLQPQRQNQQPQGQLAQRQVMQNSVQG
HSGAMMMGGPAPPQNPQNPGDLQTSSGQNQPHP
GN

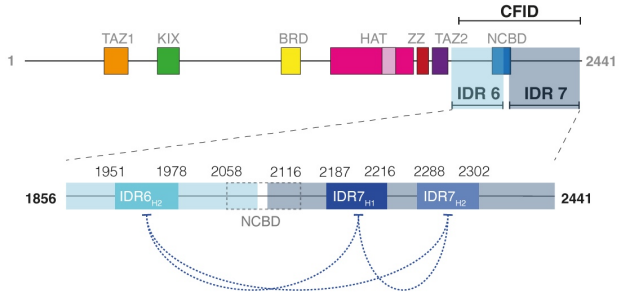
D



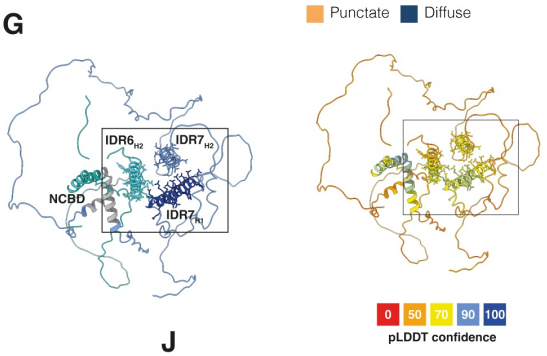
E



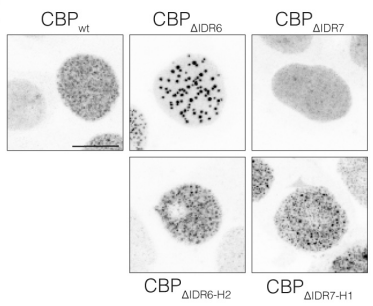
F



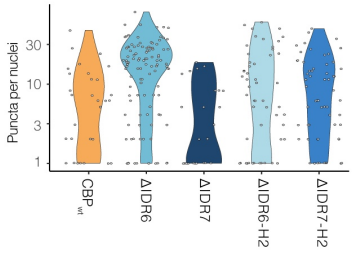
G



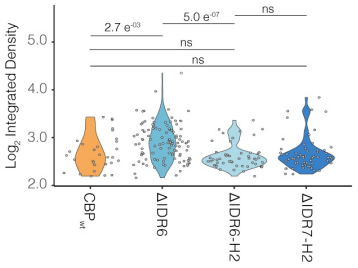
H



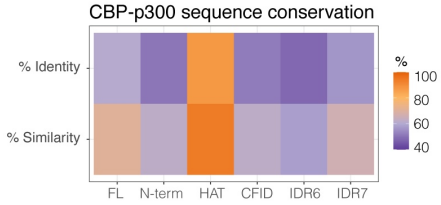
I



J



K



L

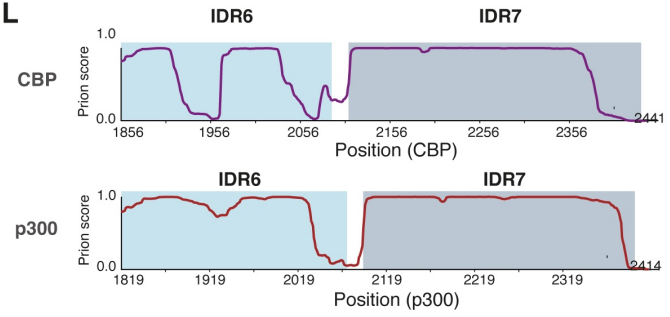


Figure S3

A) Sequence properties of CBP_{IDR6} and CBP_{IDR7} (Top panel) and the same equivalent domains following random shuffling ($CBP_{IDR6\text{-}shuffle}$ and $CBP_{IDR7\text{-}shuffle}$,bottom panel). Data shows: Fraction of Charged Residues (FCR); Net Charge Per Residue (NCPR); Linear mixing of charged residues (κ)[47]; Linear mixing of charged residues and proline (Ω)[46]. All values were calculated using LocalCIDER[43].

B-C) Randomly shuffled sequences for: B) $CBP_{IDR6\text{-}shuffle}$ and; C)$CBP_{IDR7\text{-}shuffle}$ calculated using GOOSE[48].

D) Percentage of transfected nuclei for each condition remaining after filtering.

E) Percentage of punctate nuclei CBP_{wt}, $CBP_{IDR6\text{-}shuffle}$ and $CBP_{IDR7\text{-}shuffle}$.

F) Domain structure of the CBP_{CFID} highlighting the positions of CBP_{IDR6} (cyan) and CBP_{IDR7} (light blue). NCBD and predicted helices are highlighted. Dashed lines show the interactions between α-helices predicted by AlphaFold v2.3.2 (Jumper et al., 2021).

G) AlphaFold v2.3.2 (Jumper et al., 2021) prediction of CBP's C-terminal domains. Left panel: The positions of the predicted interacting α-helices CBP_{IDR6H2} (cyan), CBP_{IDR7H1} (dark blue), CBP_{IDR7H2} (light blue), and the NCBD are highlighted; Right Panel) pLDDT confidence scores.

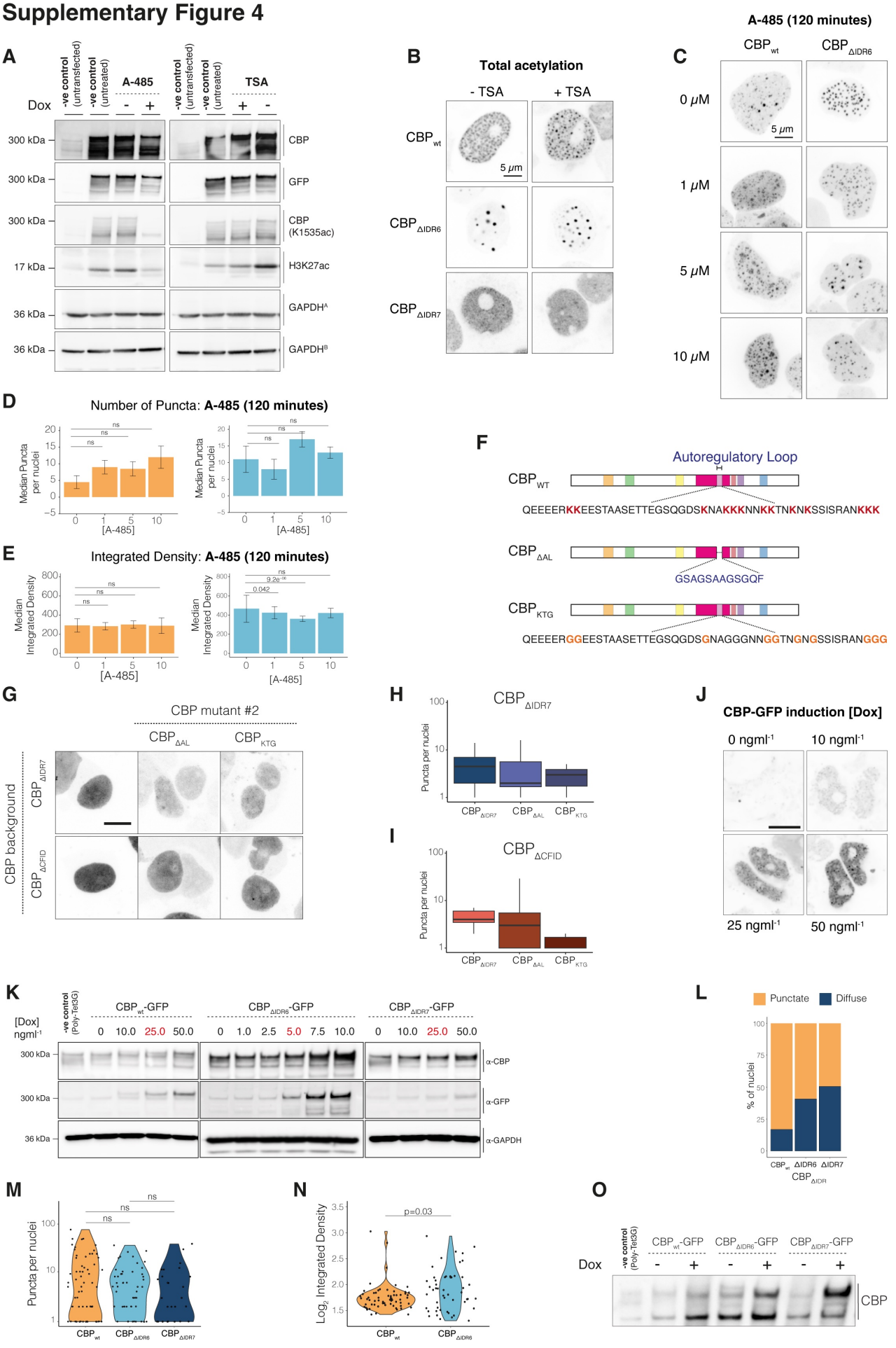
H) Fixed cell confocal microscopy comparing the behaviour of CBPwt-GFP, CBPΔIDR6-GFP and $CBP_{\Delta IDR7}$-GFP with $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP. Scale bars: 10 µm.

I-J) Comparison of condensates formed by $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. I) Number of puncta per nuclei; J) Integrated intensity density. p values were calculated using a Kruskal-Wallis test.

K) Amino acid sequence conservation between CBP and p300 from local pairwise sequence alignment using EMBOSS Matcher(Madeira et al., 2022). N-term, regions N-terminal to the HAT domain residues CBP 1-1195; p300 1-1159. (Full alignments are available in Supplementary material).

L) PLAAC analysis(Alberti et al., 2009; Lancaster et al., 2014) of CBP (purple) and p300 (Magenta) C-terminal regions. The relative positions of CBP_{IDR6} and CBP_{IDR7} are highlighted. Positions in p300 are based on local pairwise sequence alignment using EMBOSS Matcher(Madeira et al., 2022)(Full alignments are available in Supplementary material).

Supplementary Figure 4



1208 Figure S4

1209 A) Western blot of CBP, GFP, K1535ac in the AL of CBP/p300 and H3K27ac and following treatment with
1210 500nM A-485 for 36 hours[53] and 500nM Trichostatin A (TSA)[54] for 90 minutes. CBP was resolved using 3-8%
1211 Tris-Acetate PAGE (loading control: GAPDH[A]) and H3K27ac was resolved using 12% Bis-Tris PAGE
1212 (loading control: GAPDH[B]).

1213 B) Fixed cell confocal microscopy of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP after 48h treatment with
1214 YYnM Trichostatin A (TSA), DMSO or no treatment. Scale bars: 5 µm.

1215 C-E) Effect of A-485 concentration on CBP condensate behaviour: C) Fixed cell confocal microscopy of
1216 CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP after 120 minutes treatment with stated concentration of A-485 (0um =
1217 DMSO). Scale bars: 5 µm; D) Number of condensates per nuclei; E) Integrated intensity density (IID) of
1218 condensates.

1219 F) CBP_{AL} constructs showing sequences of CBP_{wt}, $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP. In $CBP_{\Delta AL}$-GFP the deleted
1220 region was replaced by a 12-residue flexible linker (blue); Lysine residues subject to acetylation are
1221 highlighted (red, bold); positions of K>G mutations are highlighted (orange, bold).

1222 G) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm

1223 H-I) Number of condensates per nucleus for $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP on a background of: H)
1224 $CBP_{\Delta IDR7}$-GFP and I) $CBP_{\Delta CFID}$-GFP.

1225 J) Inducible CBP-GFP expression: CBP_{wt}-GFP HEK293T cells were treated with 0-50ng/ml Doxycycline for
1226 48h. Confocal microscopy images. Scale bars: 10µm;

1227 K) Inducible CBP-GFP expression: western blot for CBP and GFP. The concentration of Doxycycline that
1228 induced even CBP-GFP expression is highlighted (red)

1229 L-N) Behaviour of CBP-GFP condensates following Dox induction. L) Percentage of nuclei displaying diffuse
1230 or punctate signal; M) Number of condensates per nuclei; N) Integrated intensity density of condensates
1231 formed by $CBP_{\Delta IDR}$-GFP constructs.

1232 O) Immunoprecipitated CBP-GFP used for western blot HAT assay (Figure 4O). CBP-GFP expression was
1233 induced at equivalent levels using Doxycycline as in (Figure 4M & S4K) and CBP-GFP was
1234 immunoprecipitated using GFP-Trap® Magnetic beads; CBP-GFP was resolved using 3-8% PAGE.

1235 *All p-values for IID were calculated using Kruskal-Wallis tests.*

Supplementary Figure 5

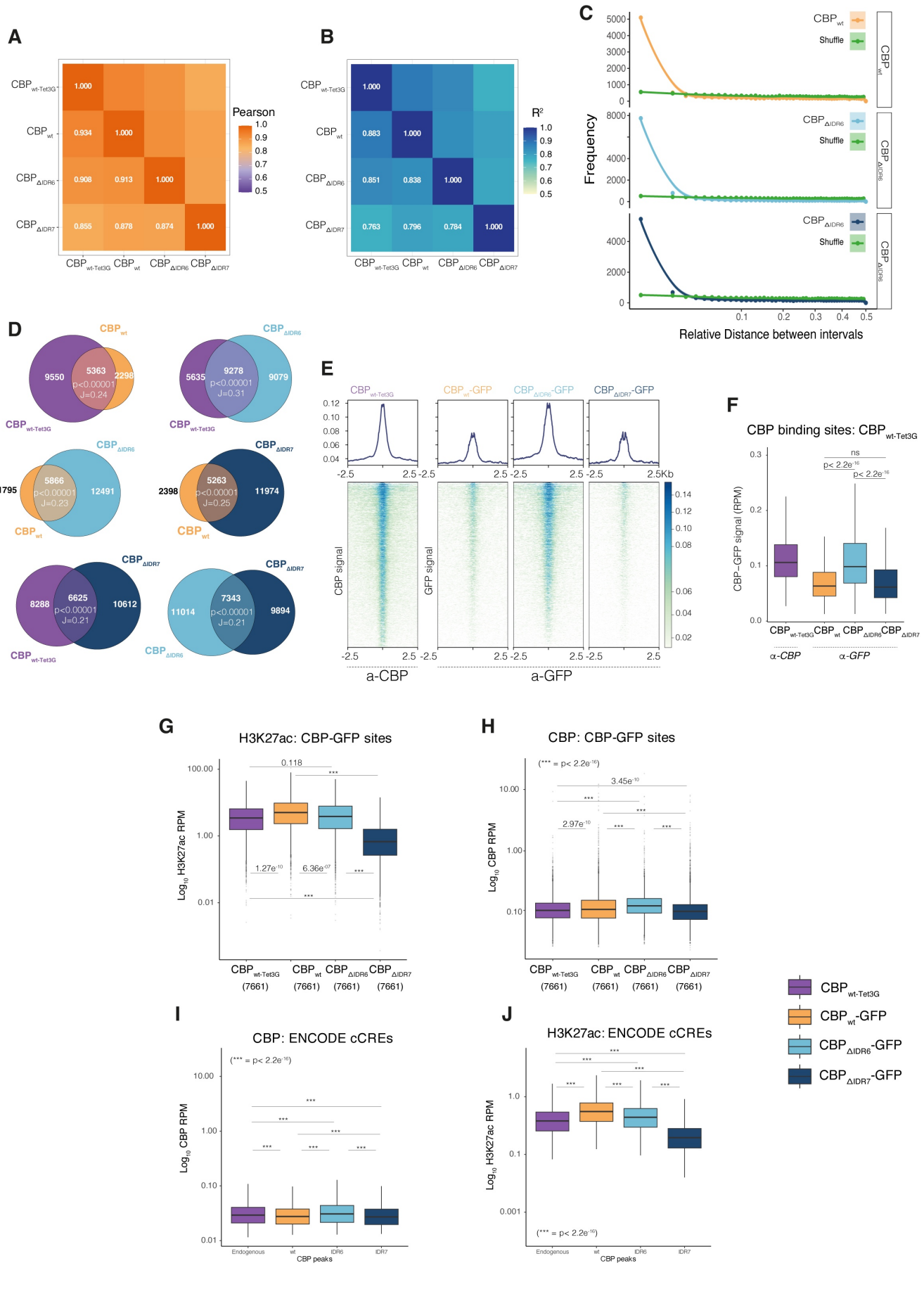


Figure S5: CBP-IDRs regulate nuclear chromatin association and H3K27ac.

A-B) Correlation between ChIPseq datasets: A) Pearson's correlation coefficient; B) R^2.

C) Relative distance plots for CBP-GFP ChIP-seq peaks (GFP ChIP). When intervals are closer than expected by chance, distributions shift towards lower relative distance values. Shuffled control regions (green) were restricted to TSS -40 kb.

D) Venn diagrams showing overlap between CBP binding sites in maternal $HEK293T_{Tet3G}$ lines (purple) and CBP-GFP cell lines; CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from permutation test with random regions restricted to TSS -40 kb; Jaccard statistic (J) measures overlap between two sets of intervals; perfect correlation = 1.

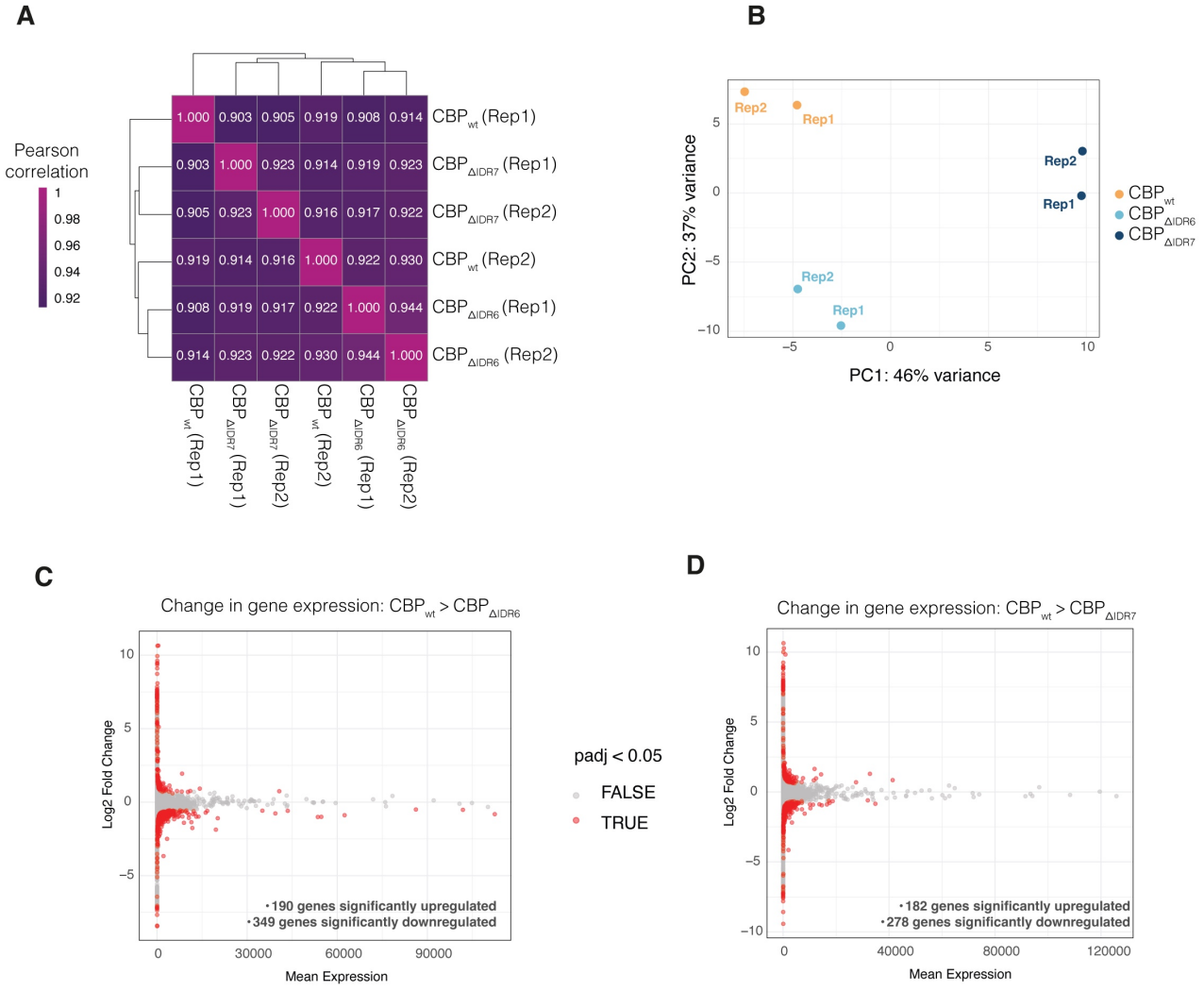
E) Heatmaps of CBP and GFP ChIP signal at 14,913 endogenous CBP_{wt} peaks from $CBP_{wt\text{-}Tet3G}$. ±2.5-kb window from center of CBP peak.

F) Reads at endogenous CBP_{wt} peaks. CBP ChIPseq reads in $HEK293T_{Tet3G}$ cells (purple); GFP ChIPseq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from Mann-Whitney U-test.

G-H) Read density enrichment at 7,661 CBP_{wt}-GFP peaks. E) H3K27ac; F) CBP-GFP. $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP sites were randomly downsampled to 7,661 peaks; p values from Mann-Whitney U test.

I-J Enrichment of I) CBP-GFP and J) H3K27ac read densities at 2,348,854 GRCh38 cCREs from ENCODE v4 [56] . p-values from Mann-Whitney U test.

Supplementary Figure 6



1295 Figure S6

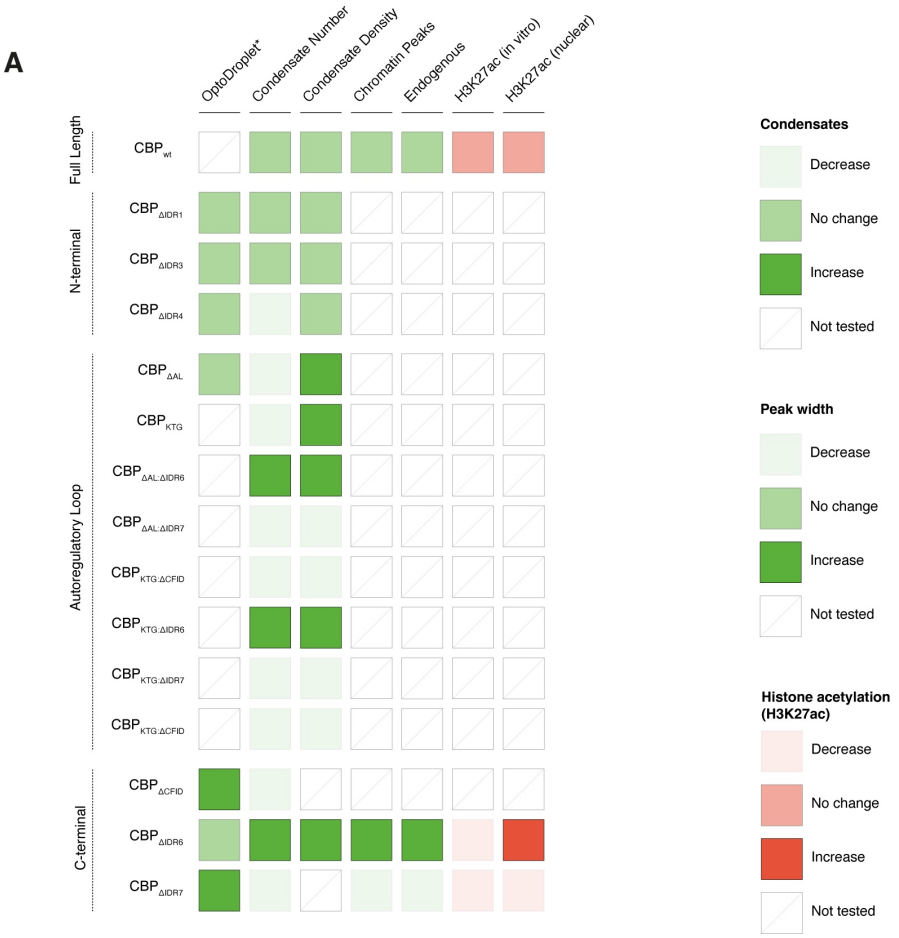
1296 A) Pearson correlation coefficient between RNA-seq datasets from inducible CBP-GFP HEK293T cell lines
1297 following induction of CBP-GFP expression using doxycycline (Figure 4M & S4K). Data shows two
1298 biological replicates (Rep1/Rep2) for each CBP-IDR deletion.

1299 B) Principle Components Analysis (PCA) of all RNAseq datasets.

1300 C-D) Gene expression differences between: C) CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP; and D) CBP_{wt}-GFP and
1301 $CBP_{\Delta IDR7}$-GFP. Genes showing significant changes in expression ($p<0.05$) are highlighted (red points).

1302

Supplementary Figure 7



1310 Figure S7

1311 B) Summary of the contribution of CBP-IDRs to CBP condensate behaviours relative to CBP_{wt}-GFP. IDRs are
1312 clustered according to whether they are: N-terminal to the HAT domain; in the Autoregulatory Loop (AL); or
1313 C-terminal to the HAT domain. All behaviours refer to behaviour of CBP following deletion of each IDR,
1314 except the OptoDroplet data which is the behaviour of the IDR in isolation. Endogenous = behaviour of
1315 endogenous CBP condensates following transfection with CBP-GFP construct.

1316

References

1. Holehouse, A.S., and Kragelund, B.B. (2023). The molecular basis for cellular function of intrinsically disordered protein regions. Nat. Rev. Mol. Cell Biol., 1–25. https://doi.org/10.1038/s41580-023-00673-0.

2. Lobel, J.H., and Ingolia, N.T. (2025). Deciphering disordered regions controlling mRNA decay in high-throughput. Nature, 1–9. https://doi.org/10.1038/s41586-025-08919-x.

3. Moses, D., Guadalupe, K., Yu, F., Flores, E., Perez, A.R., McAnelly, R., Shamoon, N.M., Kaur, G., Cuevas-Zepeda, E., Merg, A.D., et al. (2024). Structural biases in disordered proteins are prevalent in the cell. Nat. Struct. Mol. Biol. https://doi.org/10.1038/s41594-023-01148-8.

4. Cuevas-Velazquez, C.L., Vellosillo, T., Guadalupe, K., Schmidt, H.B., Yu, F., Moses, D., Brophy, J.A.N., Cosio-Acosta, D., Das, A., Wang, L., et al. (2021). Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. Nat. Commun. *12*, 5438. https://doi.org/10.1038/s41467-021-25736-8.

5. Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131–3141.

6. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. *107*, 21931–21936. https://doi.org/10.1073/pnas.1016071107.

7. Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Bühler, M., et al. (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. Cell *152*, 859–872. https://doi.org/10.1016/j.cell.2013.01.032.

8. Narita, T., Higashijima, Y., Kilic, S., Liebner, T., Walter, J., and Choudhary, C. (2023). Acetylation of histone H2B marks active enhancers and predicts CBP/p300 target genes. Nat. Genet. *55*, 679–692. https://doi.org/10.1038/s41588-023-01348-4.

9. Weinert, B.T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B.K., Schölz, C., Hamilton, W.B., Zucconi, B.E., Wang, W.W., Liu, W.R., et al. (2018). Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome. Cell *174*, 231–244.e12. https://doi.org/10.1016/j.cell.2018.04.033.

10. Boija, A., Mahat, D.B., Zare, A., Holmqvist, P.-H., Philip, P., Meyers, D.J., Cole, P.A., Lis, J.T., Stenberg, P., and Mannervik, M. (2017). CBP Regulates Recruitment and Release of Promoter-Proximal RNA Polymerase II. Mol. Cell *68*, 491–503.e5. https://doi.org/10.1016/j.molcel.2017.09.031.

11. Ma, L., Gao, Z., Wu, J., Zhong, B., Xie, Y., Huang, W., and Lin, Y. (2021). Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. Mol. Cell *81*, 1682–1697.e7. https://doi.org/10.1016/j.molcel.2021.01.031.

12. Hunt, G., Boija, A., and Mannervik, M. (2022). p300/CBP sustains Polycomb silencing by non-enzymatic functions. Mol. Cell *82*, 3580–3597.e9. https://doi.org/10.1016/j.molcel.2022.09.005.

13. Trojanowski, J., Frank, L., Rademacher, A., Mücke, N., Grigaitis, P., and Rippe, K. (2022). Transcription activation is enhanced by multivalent interactions independent of phase separation. Mol. Cell *82*, 1878–1893.e10. https://doi.org/10.1016/j.molcel.2022.04.017.

14. Gomez-Lamarca, M.J., Falo-Sanjuan, J., Stojnic, R., Abdul Rehman, S., Muresan, L., Jones, M.L., Pillidge, Z., Cerda-Moya, G., Yuan, Z., Baloul, S., et al. (2018). Activation of the notch signaling pathway in vivo elicits changes in CSL nuclear dynamics. Dev. Cell *44*, 611–623.e7. https://doi.org/10.1016/j.devcel.2018.01.020.

15. Demarest, S.J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H.J., Evans, R.M., and Wright, P.E. (2002). Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. Nature *415*, 549–553.

16. Kjaergaard, M., Andersen, L., Nielsen, L.D., and Teilum, K. (2013). A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. Biochemistry *52*, 1686–1693. https://doi.org/10.1021/bi4001062.

17. Lee, C.W., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2010). Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. Biochemistry *49*, 9964–9971. https://doi.org/10.1021/bi1012996.

18. Lin, C.H., Hare, B.J., Wagner, G., Harrison, S.C., Maniatis, T., and Fraenkel, E. (2001). A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. Mol. Cell *8*, 581–590. https://doi.org/10.1016/s1097-2765(01)00333-1.

19. Dyson, H.J., and Wright, P.E. (2016). Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. J. Biol. Chem. *291*, 6714–6722. https://doi.org/10.1074/jbc.R115.692020.

20. Papachristou, E.K., Kishore, K., Holding, A.N., Harvey, K., Roumeliotis, T.I., Chilamakuri, C.S.R., Omarjee, S., Chia, K.M., Swarbrick, A., Lim, E., et al. (2018). A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. Nat. Commun. *9*, 2311. https://doi.org/10.1038/s41467-018-04619-5.

21. Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J., et al. (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat. Struct. Mol. Biol. *11*, 308–315. https://doi.org/10.1038/nsmb740.

22. Bose, D.A., Donahue, G., Reinberg, D., Shiekhattar, R., Bonasio, R., and Berger, S.L. (2017). RNA binding to CBP stimulates histone acetylation and transcription. Cell *168*, 135–149.e22. https://doi.org/10.1016/j.cell.2016.12.020.

23. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E., and Panne, D. (2013). Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. Nat. Struct. Mol. Biol. https://doi.org/10.1038/nsmb.2642.

24. Ortega, E., Rengachari, S., Ibrahim, Z., Hoghoughi, N., Gaucher, J., Holehouse, A.S., Khochbin, S., and Panne, D. (2018). Transcription factor dimerization activates the p300 acetyltransferase. Nature *562*, 538–544. https://doi.org/10.1038/s41586-018-0621-1.

25. Hou, T.Y., and Kraus, W.L. (2022). Analysis of estrogen-regulated enhancer RNAs identifies a functional motif required for enhancer assembly and gene expression. Cell Rep. *39*, 110944. https://doi.org/10.1016/j.celrep.2022.110944.

26. Carullo, N.V.N., and Day, J.J. (2019). Genomic Enhancers in Brain Health and Disease. Genes *10*. https://doi.org/10.3390/genes10010043.

27. Thompson, P.R., Kurooka, H., Nakatani, Y., and Cole, P.A. (2001). Transcriptional coactivator protein p300. Kinetic characterization of its histone acetyltransferase activity. J. Biol. Chem. *276*, 33721–33729. https://doi.org/10.1074/jbc.M104736200.

28. Zhang, Y., Brown, K., Yu, Y., Ibrahim, Z., Zandian, M., Xuan, H., Ingersoll, S., Lee, T., Ebmeier, C.C., Liu, J., et al. (2021). Nuclear condensates of p300 formed though the structured catalytic core can act as a storage pool of p300 with reduced HAT activity. Nat. Commun. *12*, 4618. https://doi.org/10.1038/s41467-021-24950-8.

29. Perissi, V., Dasen, J.S., Kurokawa, R., Wang, Z., Korzus, E., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1999). Factor-specific modulation of CREB-binding protein acetyltransferase activity. Proc. Natl.

Acad. Sci. U. S. A. *96*, 3652–3657.

30. Ferrie, J.J., Karr, J.P., Tjian, R., and Darzacq, X. (2022). “Structure”-function relationships in eukaryotic transcription factors: The role of intrinsically disordered regions in gene regulation. Mol. Cell. https://doi.org/10.1016/j.molcel.2022.09.021.

31. LaMorte, V.J., Dyck, J.A., Ochs, R.L., and Evans, R.M. (1998). Localization of nascent RNA and CREB binding protein with the PML-containing nuclear body. Proc. Natl. Acad. Sci. U. S. A. *95*, 4991–4996. https://doi.org/10.1073/pnas.95.9.4991.

32. Girdwood, D., Bumpass, D., Vaughan, O.A., Thain, A., Anderson, L.A., Snowden, A.W., Garcia-Wilson, E., Perkins, N.D., and Hay, R.T. (2003). P300 transcriptional repression is mediated by SUMO modification. Mol. Cell *11*, 1043–1054.

33. Boisvert, F.M., Kruhlak, M.J., Box, A.K., Hendzel, M.J., and Bazett-Jones, D.P. (2001). The transcription coactivator CBP is a dynamic component of the promyelocytic leukemia nuclear body. J. Cell Biol. *152*, 1099–1106. https://doi.org/10.1083/jcb.152.5.1099.

34. Nucifora, F.C., Sasaki, M., Peters, M.F., Huang, H., Cooper, J.K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V.L., et al. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science *291*, 2423–2428.

35. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

36. Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. Cell. https://doi.org/10.1016/j.cell.2020.04.018.

37. Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim. Biophys. Acta *1804*, 996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011.

38. Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126–130.

39. Shin, Y., Berry, J., Pannucci, N., Haataja, M.P., Toettcher, J.E., and Brangwynne, C.P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. Cell *168*, 159–171.e14. https://doi.org/10.1016/j.cell.2016.11.054.

40. Luo, F., Gui, X., Zhou, H., Gu, J., Li, Y., Liu, X., Zhao, M., Li, D., Li, X., and Liu, C. (2018). Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. Nat. Struct. Mol. Biol. *25*, 341–346. https://doi.org/10.1038/s41594-018-0050-8.

41. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066–1077. https://doi.org/10.1016/j.cell.2015.07.047.

42. Park, S., Stanfield, R.L., Martinez-Yamout, M.A., Dyson, H.J., Wilson, I.A., and Wright, P.E. (2017). Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation. Proceedings of the National Academy of Sciences *114*, E5335–E5342.

43. Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O.G., and Pappu, R.V. (2017). CIDER: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. Biophys. J. *112*, 16–21. https://doi.org/10.1016/j.bpj.2016.11.3200.

44. Chong, S., Graham, T.G.W., Dugast-Darzacq, C., Dailey, G.M., Darzacq, X., and Tjian, R. (2022). Tuning

levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. Mol. Cell *82*, 2084–2097.e5. https://doi.org/10.1016/j.molcel.2022.04.007.

45. Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*, eaar2555. https://doi.org/10.1126/science.aar2555.

46. Martin, E.W., Holehouse, A.S., Grace, C.R., Hughes, A., Pappu, R.V., and Mittag, T. (2016). Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. J. Am. Chem. Soc. *138*, 15323–15335. https://doi.org/10.1021/jacs.6b10272.

47. Das, R.K., and Pappu, R.V. (2013). Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc. Natl. Acad. Sci. U. S. A. *110*, 13392–13397. https://doi.org/10.1073/pnas.1304749110.

48. Emenecker, R.J., Guadalupe, K., Shamoon, N.M., Sukenik, S., and Holehouse, A.S. (2023). Sequence-ensemble-function relationships for disordered proteins in live cells. bioRxivorg, 2023.10.29.564547. https://doi.org/10.1101/2023.10.29.564547.

49. Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell *137*, 146–158. https://doi.org/10.1016/j.cell.2009.02.044.

50. Lancaster, A.K., Nutter-Upham, A., Lindquist, S., and King, O.D. (2014). PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. Bioinformatics *30*, 2501–2502. https://doi.org/10.1093/bioinformatics/btu310.

51. Piai, A., Calçada, E.O., Tarenzi, T., Grande, A.D., Varadi, M., Tompa, P., Felli, I.C., and Pierattelli, R. (2016). Just a Flexible Linker? The Structural and Dynamic Properties of CBP-ID4 Revealed by NMR Spectroscopy. Biophys. J. *110*, 372–381. https://doi.org/10.1016/j.bpj.2015.11.3516.

52. Ibrahim, Z., Wang, T., Destaing, O., Salvi, N., Hoghoughi, N., Chabert, C., Rusu, A., Gao, J., Feletto, L., Reynoird, N., et al. (2022). Structural insights into p300 regulation and acetylation-dependent genome organisation. Nat. Commun. *13*, 7759. https://doi.org/10.1038/s41467-022-35375-2.

53. Lasko, L.M., Jakob, C.G., Edalji, R.P., Qiu, W., Montgomery, D., Digiammarino, E.L., Hansen, T.M., Risi, R.M., Frey, R., Manaves, V., et al. (2017). Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature *550*, 128–132. https://doi.org/10.1038/nature24028.

54. Yi, F., Cai, C., Ruan, B., Hao, M., Yeo, S.K., Haas, M., Yang, F., Zhang, X., and Guan, J.-L. (2023). Regulation of RB1CC1/FIP200 stability and autophagy function by CREBBP-mediated acetylation in an intrinsically disordered region. Autophagy *19*, 1662–1677. https://doi.org/10.1080/15548627.2022.2148432.

55. Favorov, A., Mularoni, L., Cope, L.M., Medvedeva, Y., Mironov, A.A., Makeev, V.J., and Wheelan, S.J. (2012). Exploring massive, genome scale datasets with the GenometriCorr package. PLoS Comput. Biol. *8*, e1002529. https://doi.org/10.1371/journal.pcbi.1002529.

56. ENCODE Project Consortium, Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature *583*, 699–710. https://doi.org/10.1038/s41586-020-2493-4.

57. He, J., Huo, X., Pei, G., Jia, Z., Yan, Y., Yu, J., Qu, H., Xie, Y., Yuan, J., Zheng, Y., et al. (2024). Dual-role transcription factors stabilize intermediate expression levels. Cell. https://doi.org/10.1016/j.cell.2024.03.023.

58. Boeynaems, S., Dorone, Y., Zhuang, Y., Shabardina, V., Huang, G., Marian, A., Kim, G., Sanyal, A., Şen, N.-E., Griffith, D., et al. (2023). Poly(A)-binding protein is an ataxin-2 chaperone that regulates biomolecular condensates. Mol. Cell. https://doi.org/10.1016/j.molcel.2023.05.025.

59. Ferrie, J.J., Karr, J.P., Graham, T.G.W., Dailey, G.M., Zhang, G., Tjian, R., and Darzacq, X. (2024). p300 is an obligate integrator of combinatorial transcription factor inputs. Mol. Cell *84*, 234–243.e4. https://doi.org/10.1016/j.molcel.2023.12.004.

60. Xu, L., Xuan, H., He, W., Zhang, L., Huang, M., Li, K., Wen, H., Xu, H., and Shi, X. (2023). TAZ2 truncation confers overactivation of p300 and cellular vulnerability to HDAC inhibition. Nat. Commun. *14*. https://doi.org/10.1038/s41467-023-41245-2.

61. Petrij, F., Giles, R.H., Dauwerse, H.G., Saris, J.J., Hennekam, R.C., Masuno, M., Tommerup, N., van Ommen, G.J., Goodman, R.H., and Peters, D.J. (1995). Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. Nature *376*, 348–351. https://doi.org/10.1038/376348a0.

62. Wang, J., Weaver, I.C.G., Gauthier-Fisher, A., Wang, H., He, L., Yeomans, J., Wondisford, F., Kaplan, D.R., and Miller, F.D. (2010). CBP histone acetyltransferase activity regulates embryonic neural differentiation in the normal and Rubinstein-Taybi syndrome brain. Dev. Cell *18*, 114–125. https://doi.org/10.1016/j.devcel.2009.10.023.

63. Haghshenas, S., Bout, H.J., Schijns, J.M., Levy, M.A., Kerkhof, J., Bhai, P., McConkey, H., Jenkins, Z.A., Williams, E.M., Halliday, B.J., et al. (2024). Menke-Hennekam syndrome; delineation of domain-specific subtypes with distinct clinical and DNA methylation profiles. HGG Adv. *5*, 100287. https://doi.org/10.1016/j.xhgg.2024.100287.

64. Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., Kasper, L.H., Lerach, S., Tang, H., Ma, J., et al. (2011). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature *471*, 189–195.

65. Murata, T., Kurokawa, R., and Krones, A. (2001). Defect of histone acetyltransferase activity of the nuclear transcriptional coactivator CBP in Rubinstein–Taybi syndrome. Hum. Mol. Genet.

66. Kalkhoven, E., Roelfsema, J.H., Teunissen, H., den Boer, A., Ariyurek, Y., Zantema, A., Breuning, M.H., Hennekam, R.C.M., and Peters, D.J.M. (2003). Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. Hum. Mol. Genet. *12*, 441–450.

67. Mensah, M.A., Niskanen, H., Magalhaes, A.P., Basu, S., Kircher, M., Sczakiel, H.L., Reiter, A.M.V., Elsner, J., Meinecke, P., Biskup, S., et al. (2023). Aberrant phase separation and nucleolar dysfunction in rare genetic diseases. Nature, 1–8. https://doi.org/10.1038/s41586-022-05682-1.

68. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

69. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

70. R Core Team (2021). R: A Language and Environment for Statistical Computing. Preprint at R Foundation for Statistical Computing.

71. Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41. https://doi.org/10.1109/83.650848.

72. San Diego, C.U. (2022). GraphPad Prism 9.0 Software.

73. Babraham bioinformatics - FastQC A quality control tool for high throughput sequence data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

74. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

75. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform.

Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

76. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

77. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

78. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207. https://doi.org/10.1093/bioinformatics/btq351.

79. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

80. Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P., and Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nat. Biotechnol. *38*, 276–278. https://doi.org/10.1038/s41587-020-0439-x.

81. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

82. Jones, D.T., and Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics *31*, 857–863. https://doi.org/10.1093/bioinformatics/btu744.

83. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590–596. https://doi.org/10.1038/s41586-021-03828-1.

84. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70–82. https://doi.org/10.1002/pro.3943.

85. Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., et al. (2022). Inference of CRISPR edits from Sanger trace data. CRISPR J. *5*, 123–130. https://doi.org/10.1089/crispr.2021.0113.